

THIS WEEK

EDITORIALS

CHRONIC FATIGUE SYNDROME
Scientists should keep looking for a cause **p.266**

WORLD VIEW Markets for carbon won't keep forests standing **p.267**



ANIMAL BEHAVIOUR
Chickens have feelings too **p.268**

A bold unifying leap

How James Clerk Maxwell made a difference 150 years ago.

What is it that makes physicists proud to be physicists? One answer lies in James Clerk Maxwell's equations. Physicists can rejoice in a historical moment of great insight, can share in the expressions of that insight that only they can understand in any depth, can bond over surviving the didactic stress that many students experience in learning to apply them and, above all, can roam freely in deploying the power thus provided for understanding the world and, on occasion, changing it.

These communal consequences of a grand discovery might arise in any scientific discipline. But Maxwell's equations are particularly characteristic of physics in the way in which they unified previously disparate laws and provided a foundational framework for vast industries — both academic and commercial.

In this issue we celebrate the first expression of those equations by Scottish physicist Maxwell in the *Philosophical Magazine* 150 years ago. There he drew together several strands of understanding about the behaviour of electricity, of magnetism, of light, and of the ways in which these fundamental aspects of nature behave in matter. As Albert Einstein remarked, "so bold was the leap" of this work that it took decades for physicists to grasp its full significance. And although it was a wonderful expression of science at its purest, it was forged in the thoroughly practical culture of intellectuals at that time (see page 289).

By the time Maxwell began this work, it was known that a pattern of electric fields could be related to a distribution of electric charges; that single magnetic charges, in contrast, do not exist; that a moving electrical charge also generated a magnetic field; and that a time-varying magnetic field generated an electrical field.

One of Maxwell's achievements was the insight that magnetic fields can be electrically generated by varying electric fields, as well as by travelling charges. In expressing this and the established rules in a set of equations, he provided a framework that revolutionized our ability to understand and apply electricity and magnetism in disciplines ranging from astronomy to biology to telecommunications.

In materials, the application of Maxwell's equations depends on an understanding of the way in which those materials affect the propagation of electric and magnetic fields. In some crystals and in Earth's atmosphere, for example, electromagnetic waves propagate differently in different directions, and may also, on entering the medium, be split into components that propagate differently. Yet all are described by Maxwell's equations. An intriguing application today is in metamaterials, which have unnatural propagation properties that enable the development of extraordinarily powerful lenses and optical cloaking (see page 292).

It is not only in materials that these equations can be applied. Empty space was also illuminated by Maxwell. His identification of the generation of magnetic fields by changing electric fields, coupled with the known generation of electric fields by changing magnetic fields, brought the insight that waves combining electricity and magnetism might travel in a self-sustaining way. In subsequent work, he concluded that light was just such a wave. He identified a single

propagation velocity in empty space of all electromagnetic radiation — known in his time as the speed of light, now also known to apply to radio, infrared and ultraviolet waves, to X-rays and γ -rays.

That speed is determined in his equations by two characteristics of the medium in which the waves propagate — the electrical permittivity and the magnetic permeability. These have particular values in

"It is for Maxwell's insights in electromagnetism that scientists in all disciplines can be most grateful."

empty space, stimulating the question (already raised by Maxwell's predecessors) of whether space itself is filled with some sort of 'luminiferous aether' required for light to propagate. It took subsequent physicists to determine that such an aether was both non-existent and unnecessary. And it took Einstein

to realise that the speed of light in empty space is not only a constant regardless of direction but also a constant no matter where the observer is and no matter how the observer is moving.

It is due to the ever-more-unifying discoveries of Maxwell's successors, and recent, often controversial, hypotheses in that spirit (see page 286), that much emphasis is today placed on that aspect of Maxwell's equations. Maxwell did other grand things in science — and this year is also the 150th anniversary of his taking the first true colour photographs. But it is for his unifying insights in electromagnetism that scientists in all disciplines can be most grateful, and physicists most proud. ■

Into ignorance

Vote to overturn an aspect of climate science marks a worrying trend in US Congress.

As *Nature* went to press, a committee of the US Congress was poised to pass legislation that would overturn a scientific finding on the dangers of global warming. The Republican-sponsored bill is intended to prevent the US Environmental Protection Agency (EPA) from regulating greenhouse-gas emissions, which the agency declared a threat to public welfare in 2009. That assessment serves as the EPA's legal basis for regulation, so repealing the 'endangerment finding' would eliminate its authority over greenhouse gases.

That this finding is scientifically sound had no bearing on the decision to push the legislation, and Republicans on the House of Representatives' energy and commerce committee have made clear their disdain for climate science. At a subcommittee hearing on 14 March, anger and distrust were directed at scientists and respected scientific societies. Misinformation was presented as fact, truth was twisted and nobody showed any inclination to listen to scientists, let alone learn from them.

It has been an embarrassing display, not just for the Republican Party but also for Congress and the US citizens it represents.

It is tempting to write all of this off as petty partisanship, a populist knee-jerk reaction to lost jobs and rising energy prices by a well-organized minority of Republican voters. After all, US polling data has consistently shown that, in general, the public accepts climate science. At a hearing last week, even Ed Whitfield (Republican, Kentucky), who chairs the subcommittee, seemed to distance himself from the rhetoric by focusing not on the science but on the economic effects of greenhouse-gas regulation. "One need not be a sceptic of global warming to be a sceptic of the EPA's regulatory agenda," said Whitfield.

Perhaps, but the legislation is fundamentally anti-science, just as the rhetoric that supports it is grounded in wilful ignorance. One lawmaker last week described scientists as "elitist" and "arrogant" creatures who hide behind "discredited" institutions. Another propagated the myth that in the 1970s the scientific community warned of an imminent ice age. Melting ice caps on Mars served to counter evidence of anthropogenic warming on Earth, and Antarctica was falsely said to be gaining ice. Several scientists were on hand — at the behest of Democrats on the subcommittee — to answer questions and clear things up, but many lawmakers weren't interested in answers, only in prejudice.

It is hard to escape the conclusion that the US Congress has entered the intellectual wilderness, a sad state of affairs in a country that has led the world in many scientific arenas for so long. Global warming is a thorny problem, and disagreement about how to deal with it is understandable. It is not always clear how to interpret data or address legitimate questions. Nor is the scientific process, or any given scientist, perfect. But to deny that there is reason to be concerned, given the decades of work by countless scientists, is irresponsible.

That this legislation is unlikely to become law doesn't make it any less dangerous. It is the attitude and ideas behind the bill that are troublesome, and they seem to be spreading. Fred Upton, the Michigan Republican who chairs the full energy and commerce committee, once endorsed climate science, but last month said — after being pinned down by a determined journalist — that he is not convinced that greenhouse-gas emissions contribute to global warming. It was yet another blow to the shrinking minority of moderate centrists in both parties.

"The US Congress has entered the intellectual wilderness."

One can only assume that Congress will find its way at some point, pressured by voters who expect more from their public servants. In the meantime, as long as it can fend off this and other attacks on the EPA, President Barack Obama's administration should push forward with its entirely reasonable regulatory programme for reducing greenhouse-gas emissions where it can, while looking for ways to work with Congress in other areas. Rising oil prices should increase interest in energy security, a co-benefit of the greenhouse-gas and fuel-efficiency standards for vehicles that were announced by the administration last year. The same advice applies to the rest of the world. Work with the United States where possible, but don't wait for a sudden change of tenor in Washington DC.

One of the scientists testifying before Whitfield's subcommittee was Christopher Field, director of the Carnegie Institution's global ecology department in Stanford, California. Field generously hoped that his testimony at last week's hearing took place "in the spirit of a genuine dialogue that is in the best interests of the country". Maybe one day that hope will be justified. ■

Cause for concern

Scientists studying diseases should be motivated by patients, but not led by them.

In the furious debate over whether chronic fatigue syndrome is linked to a mysterious retrovirus called XMRV, scientists can at least agree on one thing: the whole thing is a mess.

In the year and a half since researchers first claimed they had found the virus — which resembles viral sequences in the mouse genome — in people with the disease, nothing close to consensus has emerged. Several groups have tried to verify the results using different methods, and most have found nothing, leading them to suggest that the initial, promising experiments were simply an artefact of laboratory contamination.

But the few scientists who have found evidence to link the virus to the syndrome, also known as myalgic encephalomyelitis, are rigorously defending their data (see page 282). Led by Judy Mikovits at the Whittemore Peterson Institute for Neuro-Immune Disease in Reno, Nevada, they blame patient heterogeneity, the geographical distribution of the virus and methodological issues among the reasons for the failure to verify a link.

Clarity is urgently needed, both by the scientists and by patients desperate for an answer. Owing to the difficulty in diagnosing a chronic disorder with diffuse symptoms, many people with the disease feel they have been marginalized by the medical community — told that there is no help for them, or even that they are imagining their symptoms. It's not hard to see why patients and advocacy groups have come to Mikovits' defence. XMRV provides legitimacy, and the hope of a treatment.

Given the long and convoluted history of infectious agents proposed and later disproved as a cause for chronic fatigue, this debate could continue for some time. Among the researchers searching for a definitive

answer are several high-profile groups commissioned by health officials in the United States; they also aim to determine whether XMRV is floating around in the blood supply and being passed from person to person. There is a lot at stake. Hypotheses that are later discarded by science can linger on in the minds of the public: witness the supposed link between the measles, mumps and rubella (MMR) vaccine and autism, which has refused to die years after the weight of evidence should have squashed it. Patients who latch onto XMRV as the cause of chronic fatigue syndrome are likely to find it hard to let go, even if most scientists eventually conclude that the evidence is against it. Some are already being tested for the virus and starting to take potent, antiretroviral drugs if they test positive, all on the basis of an unproven hypothesis.

The challenge for scientists in this field, as in any other that involves patients, is to understand and be motivated by the plight of the patient community without letting their research be swayed by it.

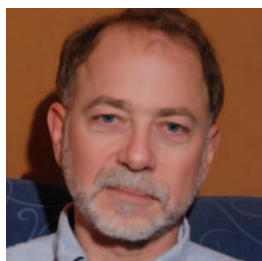
Mikovits has worked with her critics to identify and pursue the research needed to resolve the debate. She is right to engage in debate about her science and defend it where necessary. But Mikovits and her critics need to maintain an open mind.

Scientists are taking the risk of XMRV seriously, as they should. The National Institutes of Health and others examining the link should see their studies through. A thorough, well-funded effort to get to the bottom of the situation should help to ensure that time, money and the careers of young scientists are spent effectively. But if the association is not borne out by these studies (and the preponderance of evidence thus far suggests that it may not be), will the spotlight on chronic fatigue syndrome endure or dim?

Many scientists who have waded into the debate are experts in viral disease, and they say they are unlikely to pursue work on chronic fatigue. This is a pity but, given the pile-up of criticism that was Mikovits' reward for diligently following up her data, perhaps no surprise. Several lines of evidence suggest that chronic fatigue has an infectious origin. It is to be hoped that one day scientists will manage to pin it down once and for all. ■

➤ NATURE.COM
To comment online,
click on Editorials at:
go.nature.com/xbhunq

J. MANN



Cash alone will not slow forest carbon emissions

To succeed, the REDD initiative needs a dose of 'GREEN' to restore degraded forests and help boost economic development, argues Andy White.

For more than three years, the world has been trying to implement a practical way to tackle greenhouse-gas emissions from forests. Known as Reducing Emissions from Deforestation and Forest Degradation (REDD), the initiative was launched at the United Nations Climate Change Conference in Bali in 2007 with great hope and no small measure of hype. Progress on REDD was hailed as a major success at the UN climate talks in Cancún, Mexico, late last year.

The ensuing experience has convinced me that it is time for some mid-course corrections. What's more, a host of new studies shows that REDD will fail unless governments give higher priority to restoring degraded forests and promoting community conservation and enterprises in forest areas.

The goal of REDD was to reduce emissions from deforestation and forest degradation, recognized in 2007 as some 20% of global carbon emissions. It was boosted by the *Stern Review of the Economics of Climate Change*, which stated in 2006 that curbing deforestation would be both cheap and fast compared with other options to reduce emissions. By 2008, governments agreed that more than US\$20 billion a year would be needed to convince forest owners in developing countries to leave their trees standing. REDD quickly focused on luring private capital and setting up carbon markets, and drew the support of industries in developed countries that were keen to offset emissions. Developed countries earmarked \$4.5 billion to help governments in developing countries to prepare to trade forest carbon.

The carbon market has proved more difficult and expensive to develop than many expected. Property rights in rural and forested areas of many developing countries are unclear and contested, and there is little agreement on who owns the land, much less on who owns the forest or the carbon. Serious scientific challenges hamper efforts to measure carbon and monitor changes in land use. And a new study by the US-based Munden Project, which specializes in designing commodity markets, shows that a global forest carbon market could easily be manipulated and would not reduce deforestation.

Yet as the world struggles to develop REDD, national and international data show that deforestation is declining. Recently released UN figures show that, from 1990 to 2010, the net forest area increased in 58 countries that have more than 200,000 hectares of total forest, and was holding steady in another 18. Strikingly, 62% of those 76 nations are classed by the International Monetary Fund as emerging or developing countries, 8% as heavily indebted poor countries and 30% as advanced economies. Countries both rich and poor were thus protecting and restoring forests long before REDD. The estimated contribution of forest degradation to overall carbon

emissions is now as low as 8%, according to a recent report by Winrock International, a non-profit organization based in Little Rock, Arkansas.

These data reveal the central role of government choice. Studies compiled by the Rights and Resources Initiative (RRI) in Washington DC show that governments, not local people, are the primary drivers of global deforestation. Government policy and investment, generally encouraged by global demand for food, energy and wood fibre, sets off a chain reaction that leads to forest degradation or destruction. This is no surprise, given that governments still claim ownership over some 70% of tropical forests globally.

If government policy is behind the majority of deforestation, it is hard to see how cash payments through REDD would bring change. In Indonesia, the government-sponsored palm-oil industry generated more than \$12 billion in government revenues alone in 2010 — much more than the \$1 billion offered by Norway to establish REDD. Even if the carbon market works, REDD cannot compete.

The focus of REDD on finance has blinded us to other approaches to reducing forest emissions. Research shows that where indigenous peoples and forest communities have their rights recognized, they are far better forest stewards than are governments. A study by the RRI, commissioned by the World Bank, shows that the cost per hectare of recognizing rights is orders of magnitude less than the estimated costs of REDD. South Korea, China, Vietnam and Nepal have increased their forests in recent years, and an RRI study of global forest restoration shows that they have three things in

common: sustained political commitment to reforestation, reforms to support local property rights and forest management, and local economic development.

Poverty is rife in forest areas, and both local people and their governments need employment, energy and economic growth. Worldwide, there are more than 1 billion hectares of degraded forest — land that, if restored, could produce more food, wood and bioenergy, and reduce everyone's vulnerability to climate change. In the face of growing global demand for all commodities, stopping deforestation without planting trees and creating jobs just shifts the destruction around and does not relieve the many other pressures on people and forests.

So to fix REDD, we need to focus on policies to support communities, and not markets for carbon. Alongside it, we need more efforts to restore degraded forests, increase employment and produce energy in rural areas. We at the RRI call this GREEN — Growing Restoration, Employment and Energy, Now. GREEN is not just an optional complement to REDD, it is crucial for its eventual success. ■

Andy White is coordinator of the Rights and Resources Initiative in Washington DC. e-mail: awhite@rightsandresources.org

**GOVERNMENTS,
NOT LOCAL PEOPLE,
ARE THE PRIMARY
DRIVERS OF
DEFORESTATION.**

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/lberru

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

GEOPHYSICS

Faster ice melt, higher sea levels

Ice loss on Greenland and Antarctica is accelerating at three times the rate of mountain ice loss. If it continues, this melting will dominate sea-level rise this century.

Eric Rignot at the University of California, Irvine, and his colleagues compared calculations based on 18 years' worth of data on climate and ice discharge with 8 years' worth of data from the Gravity Recovery and Climate Experiment, which uses satellite measurements to assess ice mass. The authors estimate that the rate of loss is increasing by around 36.3 gigatonnes of ice a year, with a cumulative loss of 475 gigatonnes in 2006.

At current rates, melting ice sheets, mountain glaciers and ice caps around the world, as well as the thermal expansion of the oceans, could cause sea levels to rise by up to 32 centimetres by mid-century. *Geophys. Res. Lett.* doi:10.1029/2011GL046583 (2011)

ANIMAL BEHAVIOUR

Chickens feel for each other

Rats and mice show changes in behaviour when they see close relatives in distress. It seems that chickens might

also display signs of empathy — an ability to share another's emotional state. If farm animals empathize more widely with their fellows, farmers may need to take extra measures to limit stress to animals during handling, transportation and slaughter.

Joanne Edgar at the University of Bristol, UK, and her co-workers placed individual hens and their chicks in boxes, separating hen and chicks with a clear plastic sheet, and

puffed the chicks with bursts of air. In response, the mothers' behaviour and physiology changed, with greater clucking and an increased heart beat.

Proc. R. Soc. B doi:10.1098/rspb.2010.2701 (2011)

NEUROSCIENCE

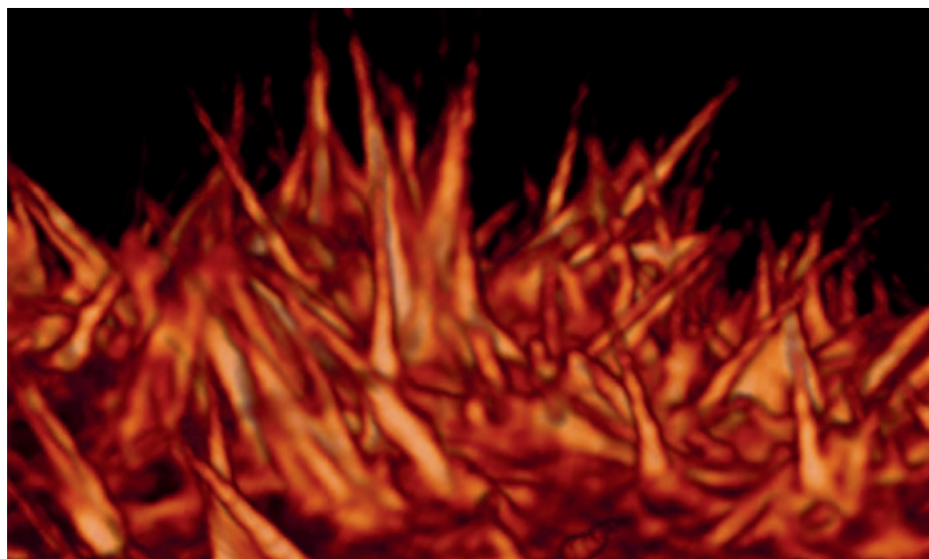
Cell support for memory

Star-shaped cells in the brain called astrocytes are thought to provide biochemical support to neurons and may even be involved in neuronal

communication. Now these cells have been found to play a part in the formation of long-term memories — by providing their neighbouring neurons with lactate.

Cristina Alberini at Mount Sinai School of Medicine in New York, Pierre Magistretti at the Swiss Federal Institute of Technology in Lausanne and their colleagues trained rats to avoid a part of a cage where they were likely to receive a mild electric shock.

When the researchers dripped a drug that blocks lactate formation in



BETZIG LAB, JANELIA FARM

CELL BIOLOGY

Seeing cells with sheets of light

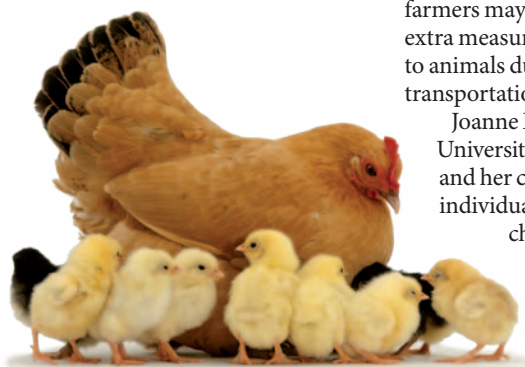
Illumination with a thin sheet of light can be used to generate high-resolution, three-dimensional movies of living cells.

The technique, called plane-illumination microscopy, had been applied to multicellular specimens, but the light sheets were too thick to capture high-resolution images in single cells. Eric Betzig of the Howard Hughes Medical Institute's Janelia Farm Research Campus in Ashburn, Virginia, and his colleagues generated thinner sheets of light using 'Bessel beams' — a

special class of narrow, non-diffracting light beams.

The team used the new microscope to produce three-dimensional images of fluorescently labelled subcellular features from stacks of planar images captured at almost 200 planes per second. The stacks were then assembled into movies showing the dynamics of certain features, such as tiny projections called filopodia (**pictured**), at a resolution of 0.3 micrometres.

Nature Methods doi:10.1038/nmeth.1586 (2011)



J. BURTON/NATURE PICTURE LIBRARY

astrocytes directly onto the rat's brains, or when they blocked the expression of lactate transporters in the hippocampus, the rats were unable to recall which part of the cage to avoid.

Cell 144, 810–823 (2011)

ASTRONOMY

New-looking old galaxies

Using a combination of space and ground-based telescopes, astronomers have spotted the most distant, and hence earliest, cluster of galaxies ever seen. The cluster dates to a time when the Universe was just a quarter of its current age of 13.7 billion years, yet, surprisingly, it looks more like nearby, modern galaxy clusters than the star-forming proto-clusters found so far at the same epoch.

Raphael Gobat of the Laboratory for Astrophysics Instrumentation and Modelling in Gif-sur-Yvette, France, and his colleagues say that more observations should show whether the cluster is a fluke, or whether theories of cluster formation need to be revised.

Astron. Astrophys. 526, A133 (2011)

MICROBIOLOGY

Stopping the cellular pump

The bacterium responsible for tuberculosis, *Mycobacterium tuberculosis*, is notorious for its tolerance to antibiotics. It enters a dormant phase, which, it was assumed, makes it less sensitive to drugs that target dividing bacteria. But it seems that even replicating mycobacteria can fend off antibiotics.

Within days of infecting zebrafish larvae with *Mycobacterium marinum*, Lalita Ramakrishnan at the University of Washington in Seattle and her colleagues found dividing drug-tolerant bacteria inside macrophages, a type of immune cell. In this environment, the mycobacteria synthesized molecular 'pumps'

that expel drugs from the bacterial cell and promote its replication.

When the team treated cultured human macrophages infected with *M. tuberculosis* with verapamil, a pump inhibitor, drug tolerance decreased, suggesting that verapamil-like drugs might be able to shorten the current six-month-minimum treatment regimen for tuberculosis.

Cell doi:10.1016/j.cell.2011.02.022 (2011)

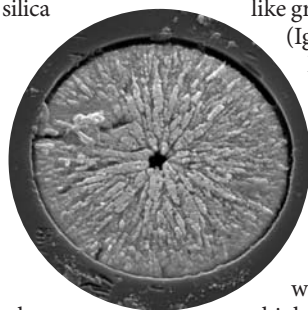
MATERIALS

Fine-tuning optical fibres

Light-transmitting fibres in fibre-optic devices are mostly made of glasses such as silica. The fabrication of crystalline compound semiconductor fibres made of zinc selenide (ZnSe), which have a wider range of desirable properties than previous fibres, opens up the possibility of developing fibre-based lasers and amplifiers for photonic applications.

John Badding at Pennsylvania State University in University Park and his collaborators deposited Zn and Se inside a hollow silica fibre (pictured) using a technique for high-pressure chemical vapour deposition. The micrometre-sized fibres had low loss of light transmission in the mid-infrared region. Their optical characteristics can be tuned by changing the composition of the material, the authors say.

Adv. Mater. doi:10.1002/adma.201003214 (2011)



NEUROSCIENCE

Brain fluid spurs stem cells

The cerebrospinal fluid (CSF) that bathes the brain and spinal cord seems to be more than just a cushion against shock. Key molecules in the CSF signal

neural stem cells to proliferate, and may be involved in a type of brain cancer.

Christopher Walsh at the Children's Hospital Boston in Massachusetts and his colleagues found that membrane proteins on the surface of neural stem cells facing the CSF bind to proteins, including insulin-like growth factor 2

(Igf2), in the fluid. Blocking this binding in mice resulted in animals with small brains. Patients with a type of brain tumour were found to have higher than normal levels of Igf2 in their CSF.

Neuron 69, 893–905 (2011)

IMMUNOLOGY

Feedback loop in lupus

The second most common human autoimmune disease, systemic lupus erythematosus (SLE), is marked by overproduction of antibodies against the body's nucleic acids. Two studies show that white blood cells called neutrophils

from patients with lupus release a network of DNA-rich fibres called neutrophil extracellular traps (NETs) that may sustain the disease.

Virginia Pascual at the Baylor Institute for Immunology Research in Dallas, Texas, and her colleagues found that neutrophils from children with SLE die *in vivo* at higher rates than those from healthy individuals. When exposed to antibodies found in the blood of patients with SLE, dying neutrophils produce NETs that activate cells called plasmacytoid dendritic cells. These produce high levels of interferon- α — which triggers immune responses, thus priming more neutrophils to release more NETs.

Michel Gilliet at the University of Texas, Houston, and his colleagues showed that people with lupus also generate antibodies against antimicrobial peptides carried by NETs, suggesting that the NETs trigger additional autoantibody production.

Sci. Trans. Med. 3, 73ra19; 73ra20 (2011)

► **NATURE.COM**

For the latest research published by Nature visit:

www.nature.com/latestresearch

COMMUNITY CHOICE

The most viewed papers in science

GENETICS

New proteins from neighbours

★ **HIGHLY READ**
on www.plosgenetics.org the weeks of 21 and 28 February

Bacteria are highly adaptive, constantly creating new proteins. Most of the genes for these proteins have been acquired from other organisms, and did not arise from the duplication and mutation of the microbes' own genes as previously thought.

Todd Treangen and Eduardo Rocha at the Pasteur Institute in Paris determined the source of more than 3,000 families of proteins in 110 bacteria belonging to 8 different groups. Depending on the group, between 88% and 98% of the new genes in the species studied came from other bacteria — through a process known as horizontal gene transfer.

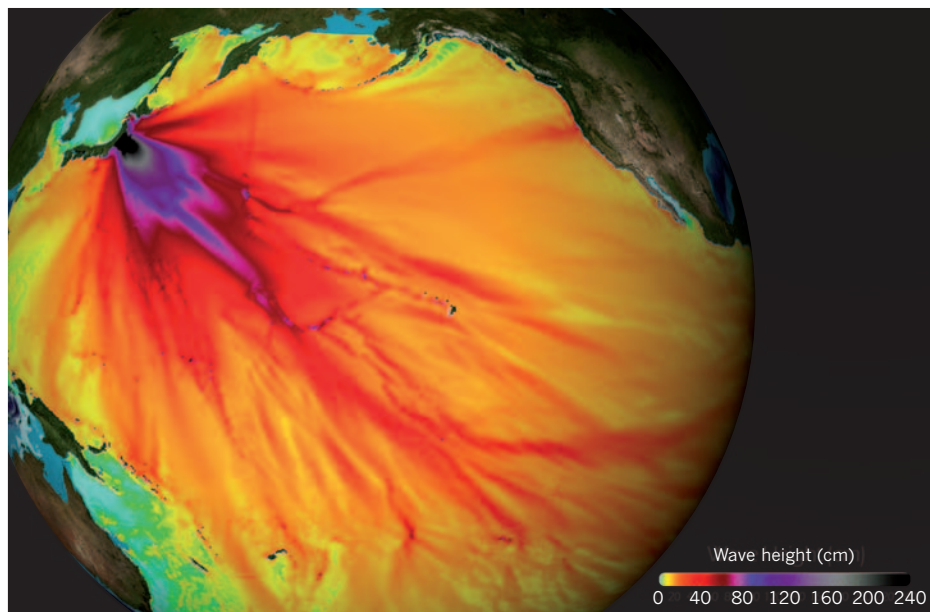
New proteins resulting from gene transfer evolve faster and persist longer than those encoded by duplicated genes, the authors show.

PLoS Genet. 7, e1001284 (2011)

SEVEN DAYS

The news in brief

NOAA



Quake causes waves of destruction

Most of the devastation seen after the magnitude-9.0 earthquake that struck Japan on 11 March was caused by the subsequent tsunami wave, which has left thousands dead. The tsunami was the largest ever measured in the open ocean by the US National Oceanic and Atmospheric Administration (which tracked its spread across the Pacific Ocean, pictured). According to local reports, the tsunami reached more than 10 metres in height when it hit

Japan's Sanriku coast barely 30 minutes after the quake. Along flat coasts, it spread hundreds of metres inland, with waves high enough to flow over concrete sea walls supposed to protect infrastructure. Thanks to the wave's long travelling time across the Pacific Ocean, and a sophisticated early-warning system, it caused only minor damage elsewhere around the Pacific. For details of the earthquake and concern over nuclear reactors, see page 273.

LARGEST EARTHQUAKES SINCE 1900

9.5 Valdivia, Chile,
22 May 1960

9.2 Prince William
Sound, Alaska,
28 March 1964

9.1 Sumatra,
Indonesia,
26 Dec 2004

9.0 Kamchatka,
former USSR,
4 Nov 1952

9.0 Sendai, Japan,
11 March 2011

8.8 Offshore Maule,
Chile,
27 Feb 2010

SOURCE: USGS

POLICY

Flu review

An independent review panel has both criticized and praised the World Health Organization (WHO) for its response to the 2009 H1N1 influenza pandemic. The 25-member panel released a preview of its findings on 10 March. The WHO gave no consistent depiction of the severity of the outbreak, and didn't dispel confusion about its definition of a pandemic, the panel found. On the other hand, the panel defended the agency against allegations by

critics that industry might have influenced WHO decisions. And it praised the WHO's leadership for its overall response. See go.nature.com/sijsov for more.

Stem-cell ethics

Procedures involving established human embryonic-stem-cell lines are not patentable, the Court of Justice of the European Communities ruled on 10 March. The judgement, on ethical grounds, was a preliminary opinion, but it surprised researchers,

who were expecting a less conservative stance. See page 280 for more.

Single EU patent

Ministers representing 25 of the 27 countries in the European Union have agreed to create a single European patent system, expected to be introduced in 2012. Currently, patents must be registered in individual countries and translated into each European language — an expensive process. Ministers exploited an 'enhanced cooperation' provision to allow progress while excluding two countries

from the agreement: Italy and Spain, who object to a transnational system that would use just English, French and German. Two days before the 10 March agreement, the European Court of Justice declared that a system of patent litigation at the European level would be unconstitutional, although this would not affect a common patent-registration system.

African lab quality

A meeting in Addis Ababa this week marked the official launch of the African Society for Laboratory Medicine, an

organization that hopes to strengthen Africa's laboratory standards and research capacity. The US President's Emergency Plan for AIDS Relief (PEPFAR) contributed US\$4.2 million to create the society, which will be based in Ethiopia. Its formation follows years of related projects with international partners such as the World Health Organization, which in 2009 launched a laboratory accreditation system to raise standards of disease diagnosis in the continent. See go.nature.com/vuehq5 for more.

PEOPLE

Vaccine move

The Global Alliance for Vaccines and Immunization (GAVI) announced on 8 March that Seth Berkley is to be its next chief executive, replacing Julian Lob-Levyt, who stepped down last October. Berkley, an epidemiologist, is giving up his role as president of the International AIDS Vaccine Initiative in New York, which he launched in 1996. He will take up the post at GAVI, which is based in Geneva, Switzerland, in August.

New JAMA head

Howard Bauchner (pictured) is to become the 16th editor-in-chief of the *Journal of the American Medical Association*



(*JAMA*), the association said on 10 March. Bauchner, a paediatrics researcher and physician currently based at the Boston University School of Medicine in Massachusetts, replaces Catherine DeAngelis, who announced last year that she would step down at the end of June 2011 after ten years in charge (see *Nature* 467, 137; 2010).

Astrobiology loss

A well-known research centre for astrobiology has lost its connection to its UK university, and will be run as a private company. Chandra Wickramasinghe, who has headed the Cardiff Centre for Astrobiology since it was founded in 2000, last week lost his appeal against Cardiff University's 2010 decision to close its astrobiology department for financial reasons (see

go.nature.com/5rbb5g). Wickramasinghe — whose work with astronomer Fred Hoyle pioneered the theory of panspermia, that life on Earth was seeded from outer space — says that the centre will now be privately funded, and will continue ongoing projects with other partners, such as the Russian space agency.

Launch-pad death

An investigation is under way into the death of a space-shuttle engineer at NASA's Kennedy Space Center in Cape Canaveral, Florida, on 14 March. James Vanover, who worked for the contractor United Space Alliance, fell from launch pad 39A, where the shuttle *Endeavour* is being prepared for its final flight in April. Officials at the space centre said that the incident was the first fatality at the launch pad since 1981, the year of the first shuttle flight.

Energy-agency chief

Dutch politician Maria van der Hoeven will be the next executive director of the International Energy Agency, succeeding Nobuo Tanaka of Japan, who steps down at the end of August after a four-year term. The agency announced van der Hoeven's appointment on 11 March. In the Netherlands, she has been a minister of economic affairs, and of science and education.

COMING UP

18 MARCH

NASA's MESSENGER spacecraft is due to enter orbit around Mercury. go.nature.com/om4yso

21–25 MARCH

The American Physical Society meets in Dallas, Texas. go.nature.com/ylcfz1

RESEARCH

California bounty

The University of Southern California in Los Angeles has received a US\$200-million gift, with no restrictions, to support its science, social science and humanities teaching and research, it announced on 9 March. The donation, the largest in the institution's history, was made by David Dornsife, chairman of steel-fabricating company Herrick Corporation in Stockton, California, and his wife Dana, founder of the Lazarex Cancer Foundation in Danville, California. The couple have previously funded a neuroscience imaging centre at the university.

Three parents

The Human Fertilisation and Embryology Authority in London is to review the safety and efficacy of assisted-conception methods used to avoid inheriting diseases caused by mutations in mitochondrial DNA, such as Leigh's disease. This potential fertility treatment, not yet allowed in the United Kingdom, transfers genetic material to create an embryo with DNA from three sources — the mother, father and a donor of mitochondrial DNA. The review's findings will be submitted to the Department of Health by mid-April.

► NATURE.COM

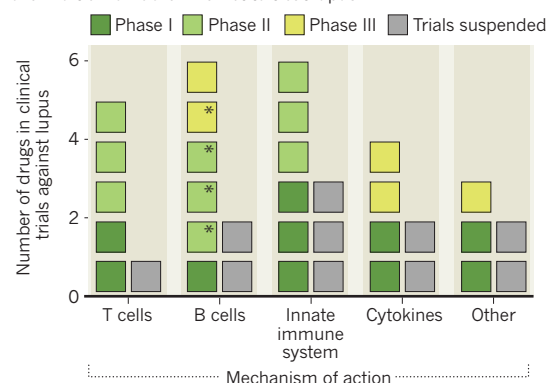
For daily news updates see: www.nature.com/news

TREND WATCH

A therapeutic antibody is the first drug in 50 years to be approved to tackle the autoimmune disease systemic lupus erythematosus. On 9 March, the US Food and Drug Administration approved Benlysta (belimumab), developed by Human Genome Sciences in Rockville, Maryland, working with GlaxoSmithKline. Other lupus therapies are hot on the trail. Some target the same protein as belimumab, hoping to reduce the body's ability to attack its own tissues. See go.nature.com/29addm for more.

A RACE TO TREAT LUPUS

Drugs that target proteins in the immune system hope to grab a share of a multibillion-dollar market to treat lupus.



NEWS IN FOCUS

EARTH SCIENCE Japan quake was a big shock to seismologists **p.274**

FUNDING Lean times could undermine US physics prowess **p.278**

EMBRYONIC STEM CELLS The ethics war takes a hidden toll **p.279**



MAXWELL'S LEGACY In search of the ultimate unification **p.286**



REUTERS/KYODO

Members of the Japanese military have already begun a clean-up operation in areas affected by radiation from the Fukushima Daiichi power station.

JAPAN

Quake sparks nuclear crisis

Explosions at a tsunami-hit plant will knock public confidence and the industry worldwide.

BY GEOFF BRUMFIEL IN LONDON AND
DAVID CYRANOSKI IN TOKYO

The magnitude-9.0 earthquake and resulting tsunami that struck Japan on 11 March have left thousands dead and many more homeless.

But the farthest-reaching consequence of the devastating natural disaster is the ongoing nuclear emergency at the Fukushima Daiichi nuclear power station. As *Nature* went to press, plant operators were struggling to cope with multiple reactors that had lost the ability to cool themselves. Three units operating at the time of the quake seem to have suffered a

partial meltdown, and some spent fuel stored at the plant has been exposed, releasing radiation.

The situation could still worsen. But scientists and engineers who have sought a revival of nuclear power for the past decade acknowledge that the accident, the worst since the Chernobyl disaster in 1986, is already a devastating blow for the technology. "It will be bad; I mean I don't think there's any point in pretending," says Robin Grimes, director of the Centre for Nuclear Engineering at Imperial College London, and an advocate of nuclear power.

The crisis began at 2:46 p.m. local time on Friday, when the massive quake struck (see 'Seismic puzzle'). Nuclear reactors at four plants

in the region immediately inserted control rods into their cores, effectively shutting them down (see 'Inside the core'). However, even after the shutdown, the decay of radioactive elements in the cores continued to pump out heat. The electric cooling pumps had lost power from the grid, but back-up diesel generators kicked into action, and ultra-pure cooling water continued to circulate through the cores.

Roughly half an hour later, however, a wall of water breached the sea wall protecting the Fukushima Daiichi plant — one of the closest to the epicentre — and struck the main back-up generators. "The problem wasn't the earthquake, but the tsunami," says Mitsuru ►

SEISMIC PUZZLE

Giant shock rattles ideas about quake behaviour

"This earthquake is a lesson in humility," says Emile Okal, a geophysicist at Northwestern University in Evanston, Illinois, who studies great earthquakes and tsunamis. Few experts had thought that the seismic zone near Sendai, Japan, was capable of producing earthquakes anywhere near as powerful as the magnitude-9.0 shock on 11 March, the largest on record in Japan. Okal and his colleagues want to understand why the event was so much stronger than many people expected — and what it means for seismic risks in Japan and elsewhere around the globe.

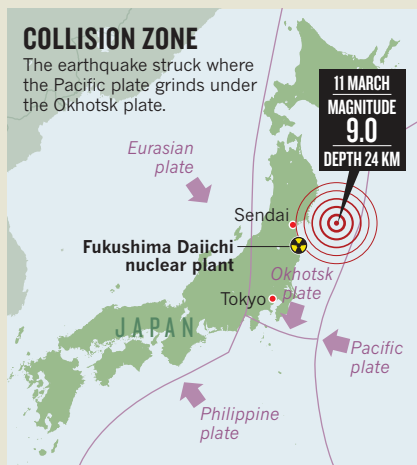
The quake happened along a seam in the planet's surface where the Pacific Ocean floor is diving beneath the tectonic plate carrying northern Japan (see 'Collision zone'). That process of subduction triggers the largest earthquakes in the world, such as the magnitude-9.5 Chilean quake in 1960 and the magnitude-9.1 Sumatran quake in 2004. But geophysicists had thought that great subduction-zone earthquakes happened only where younger oceanic crust scrapes its way into the mantle. Older crust, which is cooler and denser, was thought to slide much more readily downward, triggering smaller quakes. And the ocean crust off the northeast part of Japan, having formed about 140 million years ago, is about as old as it can get.

The history of the Sendai region seemed to support that idea. "There has been seismicity but not really great-earthquake seismicity," says Hiroo Kanamori, a seismologist at the California Institute of Technology in Pasadena. In the past few centuries, the subduction zone off the coast of Sendai has generated earthquakes of up to magnitude 8 or so, but nothing as powerful as a 9, which releases 30 times more energy.

Given that history, seismologists in Japan did not consider great earthquakes to be a threat to the Sendai area. And although that region was one of Japan's best-prepared for tsunamis, the high sea walls along much of the coast were built to stop waves far smaller than the 13–15-metre-tall giants that battered the coastline, causing most of the damage and triggering a nuclear crisis.

But some clues suggested that the Sendai region might be capable of greater violence. The giant Sumatran earthquake

► NATURE.COM
Read about the
largest earthquake
ever recorded at:
go.nature.com/hihawf



raised questions about the sea-floor-age hypothesis, because the old subducting crust there should have ruled out a shock of that size, says Okal. And recent geodetic studies across Japan showed that the Sendai region is getting squeezed, as in a vice, by pressure from the plate motions. The warping suggested that the Pacific plate was stuck rather than sliding smoothly beneath Japan, straining the crust.

The strain could only be released in earthquakes, and it builds up so quickly that the size and frequency of earthquakes seen in the recent past would not have been enough to release it, says Thomas Heaton, a geophysicist at the California Institute of Technology. It took last week's quake to do the job — and big as it was, it may not have released all the accumulated strain, says Heaton. "There's still a mystery to this place even with the 9."

If the subduction zone near Sendai can produce a great quake, then other areas with similarly old ocean crust might too, says Okal, who says that Tonga and the northeastern Caribbean are regions to look at more closely. Clues to rare, great earthquakes may be hidden there — as it seems they were in Sendai.

The last giant tsunami recorded in Sendai struck in 869. Judging from geological traces of two even older tsunami deposits, Koji Minoura, an Earth scientist at Tohoku University in Sendai, and his colleagues proposed in 2001 that giant waves visit the region about every 800–1,100 years (K. Minoura *et al.* *J. Nat. Disaster Sci.* **23**, 83–88; 2001). Because the last one came in the ninth century, the researchers wrote, "the possibility of a large tsunami striking the Sendai plain is high". **Richard Monastersky**

► Uesaka, a nuclear engineer at the University of Tokyo. The flooding is believed to have irrevocably damaged the oil tank and other key parts of the system.

With the cooling system crippled, reactor unit 1 at Daiichi was the first to heat up. Water in the core began to boil off, raising the temperature and pressure inside the massive steel containment vessel. As temperatures soared, the zirconium alloy coating the fuel apparently split or melted, reacting with the steam to produce copious amounts of hydrogen gas.

The pressure was reaching worrisome levels by the afternoon of 12 March, and Tokyo Electric Power Company (TEPCO), which runs the plant, made the bold decision to vent radioactive steam. "There was no other choice," says Masashi Goto, a former engineer at Toshiba who is an expert on the design and testing of the containment vessel. At around 3:30 p.m., shortly after TEPCO announced the venting, an explosion rocked reactor number 1. This is thought to have been caused by hydrogen from the core. The blast was powerful enough to rip the superstructure of the reactor apart, although the hardened containment vessel appears to have remained intact.

EXTREME MEASURES

Reactor operators decided to take drastic action to prevent further melting of the fuel and a possible breach of the containment vessel. At 8:20 p.m. they began to flood unit 1 with sea water, effectively ruining the reactor. As an added safety measure, they injected neutron-absorbing boric acid into the core.

Over the next two days, the two other operating reactors at Fukushima Daiichi followed a similar path. On the afternoon of 13 March, more diesel generators at the site shut down, raising fears of overheating at units 2 and 3. That evening, operators filled unit 3 with sea water and boric acid. The following morning, it too was rocked by what seems to have been a massive hydrogen explosion. Unit 2, meanwhile, reportedly lost almost all of its cooling water, and was filled with sea water on 14 March.

At 6:14 a.m. on 15 March, the pressure-suppression pool beneath the containment vessel of unit 2 was rocked by a blast. Uesaka says that, unlike previous explosions, this one was much closer to the reactor core, releasing a higher density of radioactivity and raising fears of damage to the core and the surrounding containment vessel. At around the same time, unit 4 unexpectedly caught fire. The reactor had been shut down for inspection at the time of the earthquake, and it is thought that the fire occurred when old fuel rods, stored in deep pools in the building, became exposed and overheated, releasing explosive hydrogen. Immediately after the explosion and fire, an ominous spike in radiation was detected outside the unit. Radiation monitors at the Daiichi plant briefly picked up radiation in the range of 400 millisieverts per hour, 400 times the legal

limit and the highest rate since the crisis began.

All three distressed reactors are thought to have at least partly melted down. Barring a full-scale meltdown or another fire in the spent fuel, the danger will subside with each day, as radioactive elements in the fuel decay. But most experts agree that to avert further damage, cooling will need to continue for weeks. The clean-up will take years or even decades.

The situation remained in flux as *Nature* went to press, but it is clear that the ramifications for the nuclear industry will be enormous worldwide. In Japan itself, critics were already questioning whether the nation's 54 nuclear power reactors were adequately prepared to handle earthquakes, an issue raised in 2007 after a massive quake struck the Kashiwazaki-Kariwa nuclear plant (see *Nature* 448, 392–393; 2007). Regulators in Japan will no doubt take a hard look at reactor safety in the coming weeks, which could delay the restart of even undamaged plants. Elsewhere in the region, India has decided to 'revisit' its reactor plans.

In Germany, where resistance to nuclear power is strong and vocal, the disaster immediately reopened the acrimonious debate about extending the lives of ageing reactors. In April 2002, lawmakers restricted the working life of power stations to an average of 32 years, but last December the conservative government revised the law, extending the lifespans of power stations by a further 8 years for those built before 1980 and by 14 years for those built after this date. Now the German chancellor, Angela Merkel, has put a three-month moratorium on that extension to allow time to re-evaluate safety. Seven nuclear power stations will be switched off during this period.

Switzerland has similarly put its plans to replace ageing nuclear power stations on ice. The Italian government, however, said that the disaster in Japan will not affect its plans to resume nuclear energy generation, which halted in the wake of the Chernobyl disaster. And Poland's prime minister expressed confidence in the safety of new reactor designs for the country's first nuclear power station.

The political and regulatory effects are likely to be equally far-reaching in the United States. "This disaster serves to highlight both the fragility of nuclear power plants and the potential consequences associated with a radiological release," Congressman Edward Markey (Democrat, Massachusetts) wrote in a letter to the

INSIDE THE CORE

Fukushima's boiling-water reactors are a fairly common design from the 1970s. During normal operation, nuclear fuel boils water, creating steam that can run a turbine.

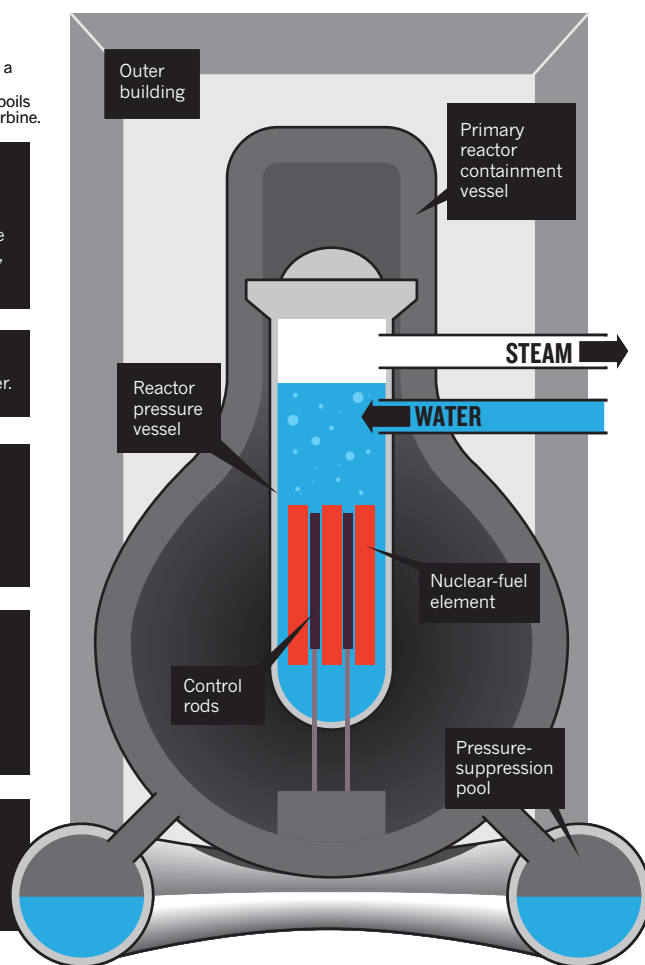
1 An earthquake strikes on 11 March; control rods at the Fukushima Daiichi power station are immediately inserted into the cores of three operating reactors, halting power generation.

2 Diesel back-up generators continue to circulate cooling water.

3 A tsunami strikes. In the hours and days that follow, cooling systems fail, causing the cores to heat up. Fuel elements may begin to deform and melt.

4 Chemical reactions between the steam and heated fuel create volatile hydrogen gas inside all three reactors. The gas escapes the core, leading to multiple explosions.

5 Operators inject sea water and boric acid into the reactor cores in order to prevent further meltdowns.



US Nuclear Regulatory Commission (NRC), demanding details about the safety of current and future reactor designs. Ellen Vancko, an analyst at the Union of Concerned Scientists, a watchdog group based in Cambridge, Massachusetts, notes that plans for a "nuclear renaissance" were already troubled by financing difficulties and the low price of fossil fuels.

But the greatest difficulty facing nuclear power may ultimately be renewed public fear, says John Large, a nuclear consultant based in London. As the accident unfolded, at least 180,000 quake-stricken people were evacuated from their homes near the Daiichi plant and a neighbouring one that also experienced cooling problems. Images of children being scanned with Geiger counters by suited safety officials evoke memories of the Chernobyl

disaster. "The real problem now, particularly in the United States and Europe, is that the public will have nothing to do with nuclear power for the next ten years," Large says.

For his part, Grimes says that he believes the event actually proves the safety of nuclear power plants. Despite being more than 30 years old, and having faced the largest earthquake ever recorded in Japan and a towering tsunami, the reactors at Fukushima Daiichi have, so far, largely contained their dangerous radioactive fuel. "Actually, it's a success," Grimes says, then adds: "Although do I think the general public will be able to see that? I think the answer is, sadly, no." ■

With additional reporting by Alison Abbott in Munich.

MORE ONLINE

TOP STORY



Nature brings you all the details of the Japanese nuclear crisis go.nature.com/hpleux

MORE NEWS

- Bacteria fight the flu go.nature.com/tzepqm
- Gravity waves draw India and Australia closer go.nature.com/zj9bzy
- Seed banks susceptible to sham samples go.nature.com/psjwbx
- Tiny PET allows scans of active animals go.nature.com/bi46y5

ON THE BLOG

Lucas Laursen completes his voyage across the Indian Ocean aboard the research ship *Hespérides* go.nature.com/ut8d8x



US free-speech law offers protection — at a price

Some fear that law to aid whistle-blowers will expose researchers to smear campaigns.

BY EUGENIE SAMUEL REICH

Daniel Klessig has had two open-heart surgeries in the past eight years, and would gladly have two more rather than face another lawsuit like the one brought against him by Meena Chandok, his former postdoctoral researcher, in 2005 after Klessig alleged that she had falsified data.

"It is the worst thing that ever happened in my life," says Klessig, who studies plant immunology at the Boyce Thompson Institute for Plant Research (BTI) in Ithaca, New York.

An institutional inquiry ultimately exonerated Chandok of misconduct, but Klessig resoundingly defeated her lawsuit. The Court of Appeals for the Second Circuit in New York ruled on 13 January that he had a "legal or moral obligation" to describe his suspicions to journal editors, funding agencies and the BTI after other scientists were unable to replicate Chandok's work. But Klessig says that years of anxiety over how

he would pay the hundreds of thousands of dollars that he might owe if he lost made the lawsuit a crippling burden.

Now, help may be at hand for whistle-blowers like Klessig, in the form of legislation soon to be introduced in the US Congress. Known as the Citizen Participation Act, the law would, its advocates say, make it easier for whistle-blowers to call out suspect research without facing the threat of a costly defamation suit, because specific provisions would enable the whistle-blowers to claim back attorneys' fees from the plaintiff in the event that they won such a case. The effort parallels a libel-reform effort in the United Kingdom, where the Ministry of Justice is moving to ease the burden of notoriously plaintiff-friendly libel laws in England and Wales (see 'Revised English law may liberate scientific discussion').

Most scientists embrace the UK reforms, which would help them to fight defamation suits brought by companies. But the proposed

US law elicits mixed reactions. Whereas some experts say that it would free whistle-blowers to speak up, others worry that it could backfire by restricting informants' ability to sue when they themselves are attacked. According to David Lewis, director of the Research Misconduct Project at the National Whistleblowers Center in Washington DC, when government agencies or companies object to a scientist's work, "their most common tactic is to charge that scientist with false allegations of misconduct".

A SLAPP IN THE FACE

Lewis has personal experience with the matter. A former research microbiologist at the US Environmental Protection Agency, he was falsely accused of scientific misconduct in 2001 by Synagro Technologies, a biosolids company in Houston, Texas, after he investigated the possibility that sewage sludge spread on farm land had harmful effects on humans. The following year he responded with a defamation lawsuit that ended in a settlement with

LEGAL REDRESS

Revised English law may liberate scientific discussion

"In just 18 months, libel has gone from being an issue no one cared about to the verge of a new defamation bill," says Simon Singh, a British science writer. Last year, Singh (**pictured**) won a widely publicized suit brought by the British Chiropractic Association over comments that he had published in the newspaper *The Guardian*. His case, along with others in which scientists were sued by companies for publicly criticizing bad science, triggered a long-awaited reform to English and Welsh libel law that is expected to be made public by the UK Ministry of Justice this week.

Singh expects the draft bill to include two measures that will redress the balance of a libel law that famously places undue burden on the defendant: he hopes it will prevent companies from having the same rights to reputation as individuals, and introduce a robust public-interest defence — something that the United States already has. "If I'm writing about a medical device or a drug



or an environmental issue, I should have additional protection," says Singh.

Sile Lane, campaigns manager for Sense

about Science in London, which is pushing for libel reform, says that the use of libel law to shut down discussion is not the only problem with the system. It also discourages people with valid cases from suing.

"At present, an ordinary person who thinks they've been defamed in a national newspaper doesn't have access to the laws," she says. Cases take too long to adjudicate, resulting in large legal bills and substantial risk for both parties, which gives an advantage to the side with most money. Steps to simplify the law and bring down the costs of defamation cases for both sides are the subject of a separate government review.

In the United States, some argue that libel reform could leave scientists who have been the subject of smear campaigns with little legal recourse. But Singh says that in England, the law has until now been so restrictive that it is hard to imagine the balance moving too far in the other direction. **E.S.R.**
Nature supports the Libel Reform Campaign.

D. CRESSEY

confidential terms. Lewis notes that the allegations haven't been repeated since he sued. (Synagro did not respond to a request for comment on the case.)

The proposed US federal law grows out of initiatives taken by several states to deal with what is known in US libel parlance as a SLAPP (strategic lawsuit against public participation). The acronym was coined in the 1980s by Penelope Canan, a sociologist now at the University of Central Florida in Orlando, and George Pring, a law professor at the University of Denver, Colorado, to describe a meritless lawsuit brought by a corporation to intimidate people who criticize government decisions favourable to those corporations. A classic example is a suit brought by a developer against residents who oppose building plans in their towns. In response to Canan and Pring's research, many states introduced statutes that would penalize parties who brought such strategic lawsuits. A focal point for the effort has been California, where the California Anti-SLAPP Project (CASP), based in Berkeley, is now helping to craft the Citizen Participation Act.

Supporters of a broad federal statute point to cases such as that of Rusi Taleyarkhan, a nuclear engineer at Purdue University in West Lafayette, Indiana, who was debarred from receiving federal funding in 2009 as a result of research-misconduct allegations brought in 2006. In 2008, Taleyarkhan sued the people who had made the allegations, including Lefteri Tsoukalas, the former head of nuclear engineering at Purdue. An Indiana court dismissed Taleyarkhan's case in December 2010, although he has since filed an appeal. Tsoukalas's defence attorney argues that the suit is a SLAPP under Indiana law, which would make Taleyarkhan responsible for both sides' legal bills, but the courts have yet to rule on this aspect of the case.

Peter Kurdock, legislative director of CASP's Public Participation Project in Washington DC, says that under the proposed federal law, whistle-blowers who could show that their comments related to a matter of public interest — such as scientific misconduct — would be able to move a case against them to a federal court where, if it

lacked merit, it could be dismissed quickly and the responsibility for paying the legal fees could be awarded to the plaintiff.



"It is the worst thing that ever happened in my life."

Daniel Klessig

Yet Canan, whom Kurdock credits with spearheading defamation reform in California, tells *Nature* that the latest legislation overreaches. "I think they're making a grave mistake," she says. Canan believes that anti-SLAPP rules ought to focus on protecting comments made by citizens to the government, which come under the last clause of the US Constitution's First Amendment: the right to petition. For scientists, that would mean extra protection for comments to a government funding agency, whereas comments to the media or journal editors would be separate issues of free speech and freedom of the press. Canan's rationale is that the government has a unique need for uninhibited input to promote good governance. "I think they're watering [the issue] down by including all free speech in it," she says.

Although an earlier version of the bill was introduced in 2009, during the last Congress, by representative Steve Cohen (Democrat, Tennessee), it never made it to a vote. Kurdock says that the alliance behind the reintroduction of a revised Act is confident that this time, it will have broad-based bipartisan support. "Libel lawsuits hit every segment of society with the explosion of free speech on the Internet," he says, "and that's why we've seen a real strong interest on Capitol Hill".

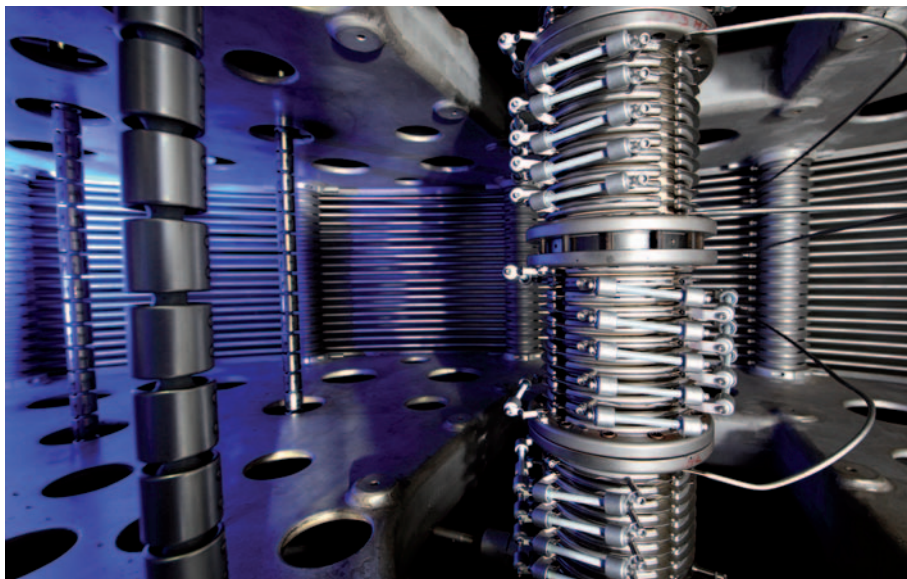
The ambiguities associated with libel are most evident in instances that pit one scientist against another. In the BTI case, for example, Klessig won his lawsuit but lost a counterclaim that Chandok's suit should be designated a SLAPP under New York state statutes. This would have increased Chandok's burden of proof and required her to pay Klessig's legal fees when she lost.

A NECESSARY MEASURE?

Klessig's lawyer, Paul Battaglia of Bond, Schoenck & King, a law firm in Syracuse, New York, says that whistle-blowers in scientific misconduct cases urgently need anti-SLAPP legislation to be broadened. "A federal statute would be beneficial because there would be a uniform standard. It would put an added burden on the researcher who brings the lawsuit and act as a disincentive for frivolous and unmeritorious lawsuits," he says.

Kurdock says that in the United States, those who win defamation suits have little to fear from anti-SLAPP statutes, because the latest law is aimed at penalizing people who bring suits with little chance of success. But as a plaintiff who has recently lost a defamation lawsuit, Chandok feels that the ruling against her is punishment enough. She had hoped that the suit would help her to get the retractions of papers that she wrote with Klessig reversed, but that now seems unlikely. "I chose to file a lawsuit for justice for myself and postdocs everywhere. Unfortunately, I feel the court has validated everything that has been done to me," she says. ■

➤ **NATURE.COM**
Read *Nature's*
Editorial arguing for
libel law reform at:
go.nature.com/ylmatk



The Holifield Radioactive Ion Beam Facility is to close, a victim of the shrinking US science budget.

RESEARCH FUNDING

US physics feels the squeeze

Obama's pro-science 2012 budget hides some bitter pills for physical scientists.

BY EUGENIE SAMUEL REICH

Joseph Bisognano sounds strained as he describes his current task: laying off 13 of the 40 staff members at the Synchrotron Radiation Center that he directs at the University of Wisconsin–Madison.

The facility produces infrared and ultraviolet photons that some 300 scientists use each year to study the structure of materials, including semiconductors and high-temperature superconductors. “Our last peer review said it would be a terrible mistake if we were closed,” says Bisognano. But the National Science Foundation (NSF) had to trim its instrument and facilities budget by 15%, so it opted to cut funding for the centre in the 2012 budget, making it likely that more lay-offs will follow.

The difficult call in Wisconsin provides a visceral glimpse of how restricted budgets are starting to squeeze some areas of the physical sciences in the United States. Although Congress and President Barack Obama have yet to agree on a final 2011 budget, stop-gap spending bills have forced the NSF and other agencies to start cutting programmes. And the president's proposal for a fiscal year 2012 budget reflects continuing pressure to cut spending.

It does request solid increases for key agencies, with large boosts in energy research, but those increases are likely to be scaled back or reversed in Congress. Even as it stands, the 2012 budget is forcing agencies to terminate mature scientific initiatives to make way for new ones.

That will undermine US supremacy in some areas of physics, warn researchers. And they complain that, in some cases, cuts have been made without consulting the scientific community. For example, the budget proposes to shutter operations at the Holifield Radioactive Ion Beam Facility at Oak Ridge National Laboratory in Tennessee, saving US\$10.3 million a year. Timothy Hallman, associate director for nuclear physics at the Department of Energy's (DOE's) Office of Science, acknowledges that the decision was not put to the department's nuclear-science advisory committee, saying the department had to act quickly.

The 2012 budget's best-known victim is the Tevatron, the particle accelerator at Fermilab in Batavia, Illinois. The DOE planned to close the facility this year as operations scaled up at the more powerful Large Hadron Collider (LHC) near Geneva in Switzerland. But the Tevatron's excellent performance prompted US researchers to request an extension, which

the DOE turned down because it would have cost \$35 million a year. Postdocs such as Elisa Pueschel of the University of Massachusetts, Amherst, saw this coming three years ago. She moved from the Tevatron to the LHC, and says that many Americans there expect to stay abroad indefinitely.

The DOE opted to close the Tevatron in part to focus on other experiments, including several aiming to capture elusive particles of dark matter or study the properties of neutrinos. Lately, though, budget concerns have hit these plans too. The DOE and the NSF have yet to reach agreement on how to fund the Deep Underground Science and Engineering Laboratory (DUSEL), to be based in Homestake, South Dakota, which would host many of these experiments. This project is estimated to cost between \$800 million and \$900 million.

The United States leads the world in developing the sophisticated beams, accelerators and detectors used in neutrino and dark-matter experiments. But it may no longer be eager to host them. “Lower-cost countries will say we'll dig the hole to put your detectors in,” says Milind Diwan of Brookhaven National Laboratory in Upton, New York, a co-spokesman for the Long Baseline Neutrino Experiment, to be based at DUSEL.

SPACE WOES

The space sciences also face tough times. Last week, a review panel recommended that NASA make a Mars sample-return mission its top priority in planetary science. But the panel assumed a budget much higher than the one NASA stands to receive, and agency officials say the mission may be unaffordable.

The James Webb Space Telescope, the replacement for the Hubble Space Telescope, is adding to NASA's budget pressure. Its cost has ballooned to \$6.5 billion — nearly half of NASA's astrophysics budget — and is squeezing mature, productive missions. For example, Suzaku, an X-ray satellite run jointly by NASA and the Japanese space agency JAXA, will lose its US support of \$2.5 million in the 2012 budget. The satellite has unrivalled resolution in the high-energy part of the X-ray spectrum. But it is not known whether JAXA will keep it going without US support.

Jon Miller, an astrophysicist at the University of Michigan in Ann Arbor who chairs the Suzaku users' group, is pessimistic about the future for US astrophysics. “High-energy astrophysics is a field where you get good return for the dollar, and it's being made to suffer,” he says. Jon Morse, astrophysics division director at NASA, says the agency sought to balance funding of existing missions with the goal of increasing the launch rate for future missions.

Rob Roser, spokesman for the Collider Detector at Fermilab experiment, says he tries to convince himself that good times will return. “My dream would be yes. But I don't know how good an imagination I have today,” he says. ■

OAK RIDGE NATL LAB, DOE

STEM CELLS

Hidden toll of embryo ethics war

Federal funds continue to be withheld for stem cells derived without destroying embryos.

BY HEIDI LEDFORD

At its heart, the ongoing legal battle to block US federal funding for research on human embryonic stem (ES) cells seeks to protect embryos.

But *Nature* has learned that in a bitter irony, the dispute seems to be holding up research on lines of human ES cells that can be derived without destroying embryos. The delay is also hampering work that researchers say could help to make adult cells a viable source of stem cells for therapies in a wide range of diseases.

In 2009, the US National Institutes of Health (NIH) unveiled draft guidelines on the human ES cell work that would be eligible for government funds. But public comments on the draft recommended a more precise definition of the eligible cells. Rather than classing human ES cells as those derived from a human embryo, as the NIH had done originally, the agency was advised to restrict the definition to cells derived from a blastocyst — an embryo of more than 100 cells. In making that change, however, the NIH inadvertently excluded a handful of lines that had been derived from a single cell — a blastomere — plucked from an eight-celled human embryo (pictured). Although deriving stem cells from a blastocyst destroys it, extracting a single blastomere — something routinely done to look for defective genes in embryos intended for *in vitro* fertilization — seems to do no harm, leaving a viable embryo that can be frozen.

Last year, the NIH proposed a further change to the ES-cell definition that would make lines derived from embryos younger than blastocysts eligible for funding (see *Nature* doi:10.1038/news.2010.85; 2010). But the agency was overtaken by the court decision in August 2010 that halted NIH support for all research on human ES cells. The court ruled that such research conflicted with prohibitions

on using government money to support work that destroys embryos. The judgement was suspended a few weeks later, allowing research

to continue until further court rulings are made (see *Nature* 470, 156–159; 2011).

The NIH has continued to approve new cell lines for federal funding under the original guidelines, but faced with so much uncertainty over the court case, the proposed changes to the definition of human ES cells have fallen by the wayside (see ‘Battle lines’). Asked when the NIH would take action on the guidelines, an agency spokesperson declined to comment.

“It’s extremely painful,” says Susan Fisher, a developmental biologist at the University of California, San Francisco, who submitted ten single-blastomere lines for NIH approval in December 2009. “We have invested so much time and effort to make these cells and now there they sit in virtual purgatory.”

In contrast, governments in other countries are funding efforts to generate single-blastomere lines for stem-cell banks. “I think this is the right approach for the field in the future,” says Carlos Simón Vallés, who heads the Valencia branch of Spain’s national stem-cell bank, and who is creating such lines. “Nobody likes to destroy embryos.”

In the United States and the United Kingdom, companies are planning to do similar work with private funding. Advanced Cell Technology (ACT), headquartered in Santa Monica, California, and Roslin Cells in Edinburgh, UK, are in discussions to establish banks of stem cells derived from blastomeres;

the embryos themselves will be frozen rather than destroyed after the procedure.

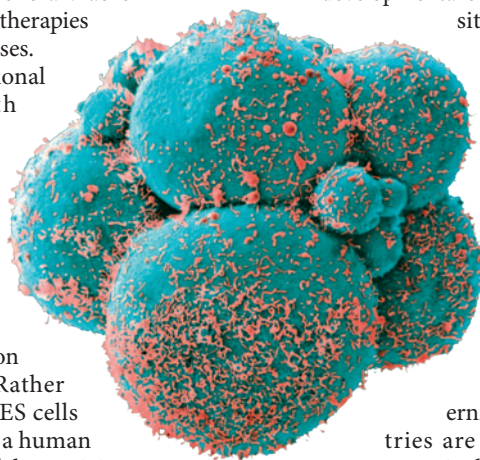
Some researchers say that the restrictions on US federal support for single-blastomere lines could hamper efforts to explore the potential of stem cells generated from adult tissue. Called induced pluripotent stem (iPS) cells, these were once heralded as a potential replacement for ES cells. But recent findings suggest that they differ in some ways from ES cells (see *Nature* 470, 13; 2011). How those differences affect pluripotency — the ability to develop into many of the body’s cell types — remains unclear. “The ultimate question for the field now is what defines pluripotency,” says Chad Cowan, a stem-cell researcher at Harvard Medical School in Boston.

Early data from Fisher’s lab — from studies funded by the California Institute of Regenerative Medicine — suggest that single-blastomere ES cells are even more malleable than those from blastocysts. “Not having federal funds used on cell lines derived from earlier embryos can stifle our opportunity to understand this pluripotent state,” says Cowan.

“Single-blastomere lines are several times more efficient at generating certain replacement cell types than are the dozens of other human embryonic stem-cell lines we’ve tested,” adds Robert Lanza, chief scientific officer at ACT, which has patented the single-blastomere technique.

For ACT, the funding restrictions also threatened to delay a clinical trial. The company had been counting on funding from the Foundation Fighting Blindness, a non-profit organization based in Columbia, Maryland, to back its trial of a therapy for Stargardt’s disease, a hereditary cause of blindness in children. But the human ES cells in the therapy were derived from a blastomere, and as the foundation draws its clinical-trial support from the federally funded National Eye Evaluation Research Network, it was barred from contributing to the costs of the trial.

“We’re tearing our hair out over here,” says Stephen Rose, chief research officer at the Foundation Fighting Blindness. “We really wanted to help fund this trial.” Eventually, ACT pulled money from its other research programmes to fund the trial, which is scheduled to begin later this year. ■



BATTLE LINES

Stem cells derived from blastomeres are stuck in regulatory limbo.

Cell lines submitted by	Number of lines	Source	Stage of embryo	NIH status
Various	86	Whole embryo	Blastocyst (> 100 cells)	Approved
George Daley, Children's Hospital, Boston	3	Whole embryo	Morula (~32 cells)	Approved, but now on hold*
Advanced Cell Technology	7	Single blastomere	Eight-cell embryo	Pending review
Susan Fisher, Univ. California, San Francisco	10	Single blastomere	Eight-cell embryo	Pending review

*Approval placed on hold because cells were derived from a pre-blastocyst embryo.

➔ NATURE.COM

For more on the stem-cell injunction, see: go.nature.com/ljonzx

EUROPEAN LAW

Europe rules against stem-cell patents

Work with human embryonic stem cells is 'contrary to ethics'.

BY ALISON ABBOTT

Stem-cell researchers in Europe are reeling after the Court of Justice of the European Communities issued an opinion last week questioning the ethics of their work and threatening to ban them from patenting procedures that involve human embryonic-stem-cell lines. Some scientists fear that the opinion could also prompt European countries to tighten their legislation on such research, or ban it altogether.

"It's the worst possible outcome," says Oliver Brüstle, director of the Institute of Reconstructive Neurobiology at the University of Bonn in Germany.

The lengthy legal debate was sparked by Brüstle's 1991 patent of a technique to generate nerve cells from established human embryonic-stem-cell lines. The environmental group Greenpeace, based in Amsterdam, challenged the patent in 2004, arguing that the destruction of human embryos involved in deriving the cell lines was "contrary to public order" and breached guidelines set out in the European Patent Convention.

In 2006, Germany's federal patent court in Munich ruled in Greenpeace's favour, so Brüstle appealed to the Federal Court of Justice in Karlsruhe. That court in turn referred the case to the European court (see *Nature* **462**, 265; 2009).

Judge Yves Bot, the case's adjudicator, concluded on 10 March that even if they do not involve the direct destruction of embryos, techniques involving human embryonic-stem-cell lines are not patentable because they are tantamount to making industrial use of human embryos, which "would be contrary to ethics and public policy".

Bot's opinion will now be considered by the 13 judges in the European court's Grand Chamber. A final decision is expected in about two months, but a spokesperson for the European court notes that few preliminary opinions are reversed. And although the final decision will not be binding on Germany's federal court, it is likely to sway the German decision.

Brüstle still hopes that his patented work will eventually lead to the generation of nerve cells that could be used to repair damage to the brain or spinal cord. But "if we are not allowed to protect our inventions in Germany, we won't be able to compete in the international market for new disease therapies", he says.



'Stop patents on life', protest Greenpeace activists.

The decision is also likely to cause trouble beyond the issue of patenting. In Europe, laws governing stem-cell research vary widely from country to country. The United Kingdom and Sweden, for instance, are relatively liberal and allow research on newly collected human embryonic stem cells that are still totipotent, meaning that they could develop into a human if they were transplanted into a womb. Others are more restrictive. Germany allows research only on imported human embryonic-stem-cell lines created before May 2007. Such cultured cell lines are merely pluripotent — capable of developing only into defined tissue types. And some countries, including Ireland, have yet to legislate on the issue.

Although Bot noted that pluripotent stem cells cannot be defined as embryos because "they are no longer capable of developing into a complete human being", he did say that the embryonic source of pluripotent cells "cannot be ignored".

As his opinion places such importance on the embryonic origin of the cell lines — irrespective of how long ago they were established — some think that the ruling is likely to encourage vacillating countries to introduce restrictive laws or complete bans on the research. Hans Schöler, a stem-cell researcher and director of the Max Planck Institute for Molecular Biomedicine in Münster, Germany, says that such countries are "quite obviously going to refer to a legal opinion at the European level for guidance". ■

S. GALLUP/GETTY IMAGES

FIGHTING FOR A CAUSE

When Judy Mikovits found links between chronic fatigue syndrome and a virus, the world took notice. Now, she's caught between the patients who believe her work and the researchers who don't.

BY EWEN CALLAWAY

On a sunny January afternoon in Santa Rosa, California, a small crowd waits patiently for Judy Mikovits to arrive. She is scheduled to deliver a talk on a mysterious virus called XMRV, which she believes underlies chronic fatigue syndrome. Although she's two hours late — held up by fog at San Francisco International Airport — not a single person has left. And when she arrives, they burst into applause.

To a rapt audience, she gives a chaotic and wide-ranging talk that explores viral sequences, cell-culture techniques and some of the criticisms that have been thrown at her since she published evidence¹ of a link between XMRV and chronic fatigue in 2009. Afterwards, Mikovits is swarmed by attendees. A middle-aged woman who spent most of the talk in a motorized scooter stands up to snap pictures of her with a digital camera. Ann Cavanagh, who has chronic fatigue and has tested positive for XMRV, says that she came in part for information and in part to show her support for Mikovits. "I just wish there were a hundred of her," Cavanagh says.

The event was "surreal", says Mikovits, a viral immunologist at the Whittemore Peterson Institute for Neuro-Immune Disease (WPI) in Reno, Nevada. She is discomfited by the attention from patients, which at times borders on adulation. But her reception among scientists has been markedly cooler. Numerous follow-up studies have found no link between the virus and the disease; no group has published a replication of her findings; and some scientists argue that XMRV is an artefact of laboratory contamination. Now, even some of Mikovits's former collaborators are having second thoughts.

Mikovits has dug in, however, attacking her critics' methods and motives. She says that their distrust of her science stems from doubts about the legitimacy of chronic fatigue syndrome itself. Chronic fatigue, also known as myalgic encephalomyelitis, affects an estimated 17 million people worldwide, but it is extremely difficult to diagnose. Many with the disorder are told that their symptoms — which include exhaustion, joint and muscle pain, cognitive issues, and heart and

respiratory problems — are psychosomatic. "I had no idea there was that much bias against this disease," Mikovits says.

The stakes are high and many are taking the risks seriously. Several countries have barred people with chronic fatigue from donating blood in case the virus spreads (see 'Something in the blood?'). And the US government has launched a US\$1.3-million study to investigate the link. Patients are already being tested for XMRV, and some are taking anti-viral drugs on the assumption that the virus causes chronic fatigue by attacking their immune defences. Many say that such action is premature, but Mikovits is steadfast. "We're not changing our course," she says.

FIRST FINDINGS

In October 2007, Mikovits attended a prostate-cancer meeting near Lake Tahoe, Nevada, where she met Robert Silverman, a virologist at the Cleveland Clinic in Ohio. Silverman co-discovered XMRV, which stands for xenotropic murine leukaemia virus-related virus². While examining human prostate tumours, he and his collaborators found genetic sequences that resemble retroviruses found in the mouse genome. Like all retroviruses, XMRV rewrites its RNA genome into DNA on infection, then slips the DNA into the genomes of host cells. Ancient remnants of such viruses litter animal genomes. But the only active retroviruses conclusively linked to human disease are HTLV-1, which causes leukaemia, and HIV.

At the meeting, Silverman was presenting research linking XMRV to deficiencies in a virus-defence pathway. Mikovits recalled that the same pathway was weakened in some patients with chronic fatigue. She wondered whether the prostate-tumour virus could also be behind chronic fatigue. After the meeting, Silverman sent Mikovits reagents to test for XMRV.

➔ NATURE.COM
Hear the author
discuss his time with
Judy Mikovits:
go.nature.com/t7ck1qx

The idea excited Mikovits, but she had other priorities. After stints in industry and at the US National Cancer Institute (NCI) in Maryland, she had recently joined the WPI to lead its research



Judy Mikovits says that she will not abandon the hypothesis that XMRV and related viruses cause chronic fatigue syndrome, despite a growing chorus of critics.

programme. The WPI was founded in 2006 by physician Daniel Peterson, an expert on chronic fatigue, and by Annette Whittemore, the wife of a well-connected Nevada businessman, whose daughter Andrea has had chronic fatigue for more than 20 years. The Whittemores spent \$5 million establishing the WPI, and several million more to support Mikovits's research, which has attracted few other grants.

At the WPI, Mikovits established a sample collection from Peterson's patients and began screening it for signs of an infection. A litany of pathogens has been linked to chronic fatigue over the years, including Epstein-Barr virus, Borna disease virus, human herpes virus 6 and HTLV-2. None panned out. Still, the disorder bears some hallmarks of an infection. Many patients report acute illness before chronic symptoms appear, and their bodies often show signs of an immune system at war. The disease can also crop up in apparent outbreaks, including one characterized by Peterson near Lake Tahoe in the 1980s.

Just before Christmas 2008, Mikovits turned her attention to Silverman's reagents. She and her postdoc, Vincent Lombardi, known as Vinny, asked a graduate student to test for XMRV DNA in white blood cells from some of the most seriously ill people being studied at the WPI.

The first try turned up just two positives out of 20. But by tweaking the conditions of the test, Mikovits says her team found XMRV in all 20. "Vinny and I looked at each other and said, 'Well, that's interesting,'" she says. They spent the next few weeks convincing themselves that they were onto something, and soon conscripted Silverman and Mikovits's former mentor at the NCI, Frank Ruscetti, to help prove that XMRV infection was behind chronic fatigue.

"We really retooled our entire programme and did nothing but focus on that," she says. They kept the effort under wraps, dubbing it 'Project

"They call me every single day. I spend so much time trying to understand the patients, to understand this disease."

X'. Even Peterson and the Whittemores weren't clued in. Mikovits says that the secrecy was necessary because her team also found XMRV in the blood of some healthy people, raising concerns about blood products. She hoped to build an airtight case because she worried that sceptical public-health officials would undermine her work.

In May 2009, the team submitted a paper to *Science* reporting the identification of XMRV genetic material in two-thirds of the 101 patients with chronic fatigue they had tested and in 3.7% of 218 healthy people. They also included data suggesting that infected white blood cells could pass the virus on to uninfected cells.

Reviewers wanted more evidence: a clear electron micrograph of virus-infected cells, proof that patients mounted an immune response to the virus, an evolutionary tree showing XMRV's relationship to other viruses and the locations where viral DNA was integrating into patient genomes. Mikovits's team went to work. "None of us took any time off, not even a weekend," she says. They resubmitted the paper in early July with everything the reviewers had asked for, except the DNA integration sites, which many scientists consider a gold standard in proving a retroviral infection.

Later that month, NCI officials who had learned about the work invited Mikovits to give a talk at a closed-door meeting with other XMRV researchers and government scientists. "When I finished speaking you could've heard a pin drop," she says. Mikovits says she thinks at least one of her manuscript's reviewers was at the meeting, because soon after, she got a call from a *Science* editor. Their paper had been accepted.

Jonathan Stoye, a retrovirologist at the MRC National Institute for Medical Research in London, wrote a commentary about the paper for *Science*³. He had never heard of Mikovits, but Frank Ruscetti's name on the paper gave him confidence, he says, and "if it were true, it was clearly very important". Stoye's co-author John Coffin, a retrovirologist at Tufts University in Boston, Massachusetts, says he was satisfied with the data and thought it was time to "let the field and public chew on them".

The BBC, US National Public Radio, *The New York Times*, *The Wall Street Journal* and dozens of other news outlets covered the research.

"Prostate cancer pathogen may be behind the disease once dubbed 'yuppie flu'," *Nature* announced on its news website the day the paper came out. Phoenix Rising, a forum for patients with chronic fatigue that has become a hub for all things XMRV, called the work a "game changer", and patients flocked to learn more about a virus that they hoped would explain their condition. But others, including Britain's leading chronic fatigue patient group, urged caution until more research buttressed the link.

The first negative findings started to arrive in January 2010 — failing to find XMRV in 186 people with chronic fatigue from the United Kingdom⁴. A month later, a team including Stoye published a paper⁵ showing no evidence of XMRV in more than 500 blood samples from patients with chronic fatigue and healthy people. One day later, the *British Medical Journal* accepted a paper reporting more negative results in Dutch patients⁶. Studies began piling up so fast that Coffin made a scorecard to show at talks. "I've lost count now," he says.

Mikovits says that the discrepancies can be explained by differences in the geographical distribution of XMRV or in the methods used.

The most common way to detect XMRV is PCR, or polymerase chain reaction, which amplifies viral DNA sequences to a level at which they can be identified. Mikovits and her team used this method to detect XMRV in some of their patients, but she contends that the most sensitive way to detect the virus is to culture patients' blood cells with a cell line in which the virus replicates more quickly. This should create more copies of the virus, making it easier to detect with PCR and other techniques. She says that none of the negative studies applied this method exactly, a fact that annoys her. "Nobody's tried to rep-ly-cate it," she says, sounding out each syllable for emphasis.

In summer 2010, some evidence emerged in Mikovits's corner. Harvey Alter, a hepatitis expert at the NIH's Clinical Center, and his team identified viruses similar to XMRV in 32 of 37 people with chronic fatigue and in 3 of 44 healthy people. They were preparing to publish their results in the *Proceedings of the National Academy of Sciences*. But scientists at the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, were about to publish a negative report. The authors delayed publication of both papers^{7,8} for several weeks to assess discrepancies. The move agitated Mikovits as well as the chronic-fatigue community, who suspected that important data were being suppressed.

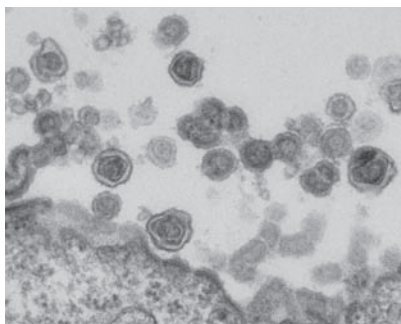
When Alter's work came out in late August⁷, Mikovits was ecstatic, and the WPI released a YouTube video of her touting it. For other researchers, however, the new paper had shortcomings. The viral sequences from Alter's paper differed from XMRV, says Greg Towers,

SOMETHING IN THE BLOOD?

Researchers and public-health officials are scrambling to determine whether blood products could be spreading the virus XMRV, which has been linked to chronic fatigue syndrome (CFS) and prostate cancer.

2009

October Judy Mikovits publishes a paper showing XMRV in two-thirds of patients with CFS and in 4% of healthy individuals¹.



December The US Department of Health and Human Services establishes a working group and the AABB (formerly the American Association of Blood Banks) organizes a task force to assess the prevalence of XMRV in blood products and the risk of its transmission.

2010

April Australia, Canada and New Zealand ban people with a history of CFS from donating blood.

June The task force recommends that patients with CFS be discouraged from donating blood in the United States.

November Britain bans patients with CFS from donating blood, officially to protect them from a decline in health.

December The working group reports discordant results of a pilot study on XMRV blood contamination, but recommends banning patients with CFS from donating blood.

The American Red Cross bans blood donations from people with a history of CFS.

2011

March At a retrovirology meeting in Boston, Massachusetts, researchers present data suggesting that XMRV is a laboratory artefact and not a human pathogen.



a retrovirologist at University College London. "He doesn't get variation, he gets a totally different virus." Towers says that mouse DNA, which is chock-full of virus sequences like those Alter's team found, probably contaminated their samples, which were collected in the 1990s. But Alter says that his team found no contamination from mouse DNA and recovered the same viral sequences from the same patients sampled a decade later.

Contamination became a dirty word for Mikovits. Just before Christmas 2010, *Retrovirology* published four papers⁹⁻¹² that highlighted laboratory contamination as a possible explanation for her findings. One showed, for example, that mouse DNA contaminates an enzyme from a commercial kit commonly used for PCR. Coffin, an author on two of the *Retrovirology* papers, urges caution against over-extrapolating. These papers do not say that contamination explains Mikovits's results, he says, just that extreme care is required to avoid it.

Towers and his colleague Paul Kellam, a virologist at the Wellcome Trust Sanger Institute near Cambridge, UK, are less charitable, however. Their study¹² showed that the XMRV sequences that Mikovits and Silverman had extracted from patients lacked the diversity expected of a retrovirus that accumulates mutations as it passes between patients. "This doesn't look like an onwardly transmissible infectious virus," says Kellam. A press release for the paper issued by the Sanger Institute put it more bluntly: "Chronic fatigue syndrome is not caused by XMRV."

Mikovits is riled when the topic turns to Towers's paper over dinner one night in Reno — "Christmas garbage", she calls it. Contamination cannot explain why her team can reproduce its results both in her lab in Reno and at Ruscetti's at the NCI, she says. Her team checks for contamination in reagents and in the cells it grows the patients' samples with. She says that her team has also collected viral sequences that will address Towers's and Kellam's criticism but that it hasn't yet been able to publish them. Meanwhile, an unpublished study of patients in Britain with chronic fatigue bears out the link to XMRV, she says. "I haven't for one second seen a piece of data that convinced me they're not infected."

Jay Levy, a virologist at the University of California, San Francisco, has a window in his closet-sized office that looks out into the laboratory where, in the 1980s, he became one of the first scientists to isolate HIV. After his discovery was scooped by other researchers, Levy turned his attention to chronic fatigue and started a long but fruitless search for an infectious cause.

Now, Levy is putting the finishing touches on what could be the most thorough response yet to Mikovits's *Science* paper, adopting the

REF. 1/AAAS

A. REEVE/SPL



Bumper stickers are just one of the supportive gifts given to the WPI.

D. CALVERT/AP

same cell-culture techniques to detect the virus and using samples from the same patients. He's done this with the help of Daniel Peterson, who left the WPI in 2010 for what Peterson says are "personal reasons". Peterson has questioned the institute's singular pursuit of XMRV, a research direction that was pursued without his consultation.

Mikovits says that she kept the XMRV work secret from Peterson over fears he would tell his patients, and left his name off the original *Science* manuscript until a reviewer questioned the omission. When asked whether that episode contributed to his departure, he says, "I was surprised at the secrecy and lack of collaboration." As for his motivation to team up with Levy: "I'm just trying to get to the truth. It's my only motive, because this is such a deserving group of patients who need to know what's going on."

Others, too, are rallying for a definitive answer. Ian Lipkin, a microbial epidemiologist at Columbia University in New York, has a reputation for getting to the bottom of mysterious disease–pathogen links. His team debunked the association between Borna disease virus and chronic fatigue, for example. Now he is spearheading the \$1.3-million effort funded by the US government. He is leaving the testing to three labs: Mikovits's at the WPI, Alter's at the NIH and the CDC. Each will receive coded samples of white blood cells and plasma from 150 patients with chronic fatigue and from 150 healthy controls. The labs will test for XMRV using their method of choice. Lipkin will crunch the data and unblind the samples.

But even if a study confirms the link to chronic fatigue, it won't be able to determine whether the virus is the cause. XMRV could, for example, be an opportunistic infection affecting those whose immune systems are already dampened by chronic fatigue. Even Mikovits can only hypothesize as to how it might cause disease.

The virus might not even exist as a natural infection. At a retrovirus conference this month in Boston, Massachusetts, Coffin and his colleague Vinay Pathak at the NCI in Frederick, Maryland, presented data showing that XMRV emerged in the 1990s, during the development of a prostate-tumour cell line called 22Rv1. Developing the line involved implanting a prostate-tumour sample into mice, retrieving cells that might divide indefinitely and repeating the process. But looking back at DNA samples taken throughout the cell-line's development showed that human cells became infected only after passing through several different mice. Importantly, XMRV's sequence seems to have come from two different mouse strains. "They just sort of snapped together like two puzzle pieces," says Coffin, an event extremely unlikely to have happened twice.

XMRV sequences retrieved from patients with prostate cancer and chronic fatigue — including some who have had chronic fatigue since the mid-1980s — are nearly identical to the virus from 22Rv1 cells. The implication, says Coffin, is that this virus, born in a laboratory, has

probably been infecting samples for more than a decade, but not people. "Although people on the blogs aren't going to believe me, I'm afraid this is by far the most reasonable explanation for how XMRV came to be," says Coffin, who hoped that the association with chronic fatigue would pan out and still thinks some pathogen other than XMRV could explain the disease.

Silverman, who no longer works with Mikovits, says that he wasn't using 22Rv1 cells when XMRV was discovered. Nonetheless, the work has rattled his confidence in XMRV's link to both prostate cancer and chronic fatigue.

Mikovits, however, is undeterred. The WPI owns a company that charges patients up to \$549 to be tested for XMRV, and Mikovits believes that patients who test positive should consult their doctors about getting antiretroviral drugs normally prescribed to those with HIV. Levy and others worry that she is overreaching. "That's scary for me. These antiretroviral drugs are not just like taking an aspirin," he says. Mikovits argues that they might be some patients' only hope. "The people who we know they're infected should have a right to get therapy," she says, "They have nothing. They have no other choice."

CONTEXT AND DEBATE

Back in her Reno laboratory two days after the talk in Santa Rosa, Mikovits examines a stack of small plastic flasks under a microscope. Some contain patient cells that she hopes will turn into cell lines and churn out XMRV. "On Wednesdays I get to take care of my cells, and that's where I'm the happiest," she says.

She has just come off the phone from a sobbing patient infected with XMRV whose symptoms had worsened. "They call me every single day," Mikovits says. "I don't do science any more. I spend so much time trying to understand the patients, to understand this disease. People have moved to Reno to be here," she says. They've left gifts: stuffed animals, and stacks of bumper stickers that say "Today's Discoveries, Tomorrow's Cures" and, more boldly, "It's the virus XMRV".

Mikovits clearly shares in the frustration of those with chronic fatigue who have been marginalized over the years and told that their disease is not real. She says that this disbelief in the disorder drives the criticism of her work. Kellam and the others say that this isn't true. They don't deny the existence of the syndrome or even the possibility of an infectious origin. "What we're trying to understand is the aetiology," Kellam says. "It's a scientific debate."

Mikovits says that she's analysed all the papers critical of her work and found flaws in each of them. Nevertheless, she's quick to endorse findings that support her work. She claims that Coffin and Pathak's study, for example, "says nothing about human infection". Yet new work presented at a different meeting that found XMRV using next-generation DNA sequencing offers "no doubt it's not contamination — that the whole story's real", she says.

Despite the growing choir of sceptics, Mikovits says that she has simply seen too many data implicating XMRV and other related viruses in chronic fatigue to change her mind. For her supporters, that steadfastness offers legitimacy and hope. "The scientists are moving forward," she announced at her talk in Santa Rosa, "and I think the politics will go away shortly." The crowd responded with vigorous applause. ■ **SEE EDITORIAL P. 266**

Ewen Callaway writes for Nature from London.

1. Lombardi, V. C. *et al. Science* **326**, 585–589 (2009).
2. Urisman, A. *et al. PLoS Pathog.* **2**, e25 (2006).
3. Coffin, J. M. & Stoye, J. P. *Science* **326**, 530–531 (2009).
4. Erlwein, O. *et al. PLoS ONE* **5**, e8519 (2010).
5. Groom, H. C. *et al. Retrovirology* **7**, 10 (2010).
6. Van Kuppeveld, F. J. *et al. Br. Med. J.* **340**, c1018 (2010).
7. Lo, S. C. *et al. Proc. Natl Acad. Sci. USA* **107**, 15874–15879 (2010).
8. Switzer, W. M. *et al. Retrovirology* **7**, 57 (2010).
9. Robinson, M. J. *et al. Retrovirology* **7**, 108 (2010).
10. Oakes, B. *et al. Retrovirology* **7**, 109 (2010).
11. Sato, E., Furuta, R. A. & Miyazawa, T. *Retrovirology* **7**, 110 (2010).
12. Hué, S. *et al. Retrovirology* **7**, 111 (2010).



Unification + 150

In 1861, James Clerk Maxwell unified electricity, magnetism and light. Experiments under way today could inch physicists closer to combining everything else.

BY M. MITCHELL WALDROP

When it happens — if it happens — don't look for Hollywood-style drama. Physicists at the Large Hadron Collider (LHC) outside Geneva in Switzerland won't suddenly gasp with astonishment, and their monitors won't flash the message, "Higgs boson detected."

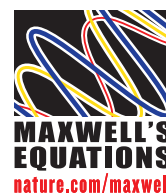
Instead, the discovery will unfold over the course of months. Computers will trawl through petabytes (10^{15} bytes) of collision data in search of a handful of distinctive events that might signal their quarry's existence, while physicists cross-check every candidate. Only when they have accumulated enough events to be sure — maybe a dozen — will they publicly proclaim the discovery of the sought-after Higgs.

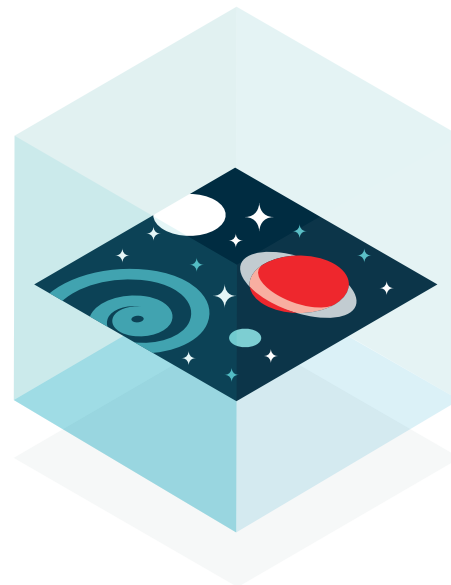
Even so, the announcement will be dramatic — and timely. Exactly 150 years ago, the

Scottish physicist James Clerk Maxwell showed that three apparently separate phenomena — electricity, magnetism and light — are different aspects of one phenomenon, today known as electromagnetism (see page 289). The Higgs discovery could take that unification a giant step further by filling in the last and most critical piece of the 'standard model', an extension of Maxwell's equations that encompasses three of the four forces of nature: electromagnetism and the weak and strong forces that act on subatomic particles. The Higgs boson is thought to interact with electrons, quarks and other fundamental particles, endowing them with mass — and thus making it possible for the standard model to describe the Universe as we know it.

This puts the standard model as it is today in the same position as Maxwell's theory before experiments demonstrated the existence of electromagnetic waves, says Frank Wilczek,

ILLUSTRATIONS BY THOMAS POROSTOCKY





a physicist at the Massachusetts Institute of Technology in Cambridge, and co-recipient of the 2004 Nobel Prize in Physics for his part in creating the model. “It looks good, lots of its predictions have been verified, but the most dramatic new thing remains to be verified.”

But even if the Higgs boson is discovered as predicted, physicists will not be satisfied. The ultimate goal is a unification theory that would reveal how all observed particles and forces are just different manifestations of a single underlying system, which can be expressed within a common mathematical framework. Such an elegant result is not possible with the standard model, which includes the strong force that binds the atomic nucleus only as an afterthought, and has nothing at all to say about gravity. The standard model also has no explanation for dark matter, an invisible substance that outweighs the ordinary matter in stars and galaxies by a factor of roughly five.

Although physicists agree that some kind of larger unification is needed, they don’t know what form that should take. For four decades, nearly as long as the standard model has existed, researchers have been speculating about ways to extend it with exotic ideas such as supersymmetry, extra dimensions and holographic space-time. “The situation is that there are a bunch of hypotheses on the table, most of them not new, with no experimental support for any of them,” says Lee Smolin, a physicist at the Perimeter Institute for Theoretical Physics in Waterloo, Canada.

“The good news,” says Smolin, “is that the experiments are finally being done.” Within a few years, thanks to the LHC and other experiments, physicists should have a much clearer idea of which theoretical notions are real and will take their place in the ultimate unification.

SUPERSYMMETRY

If the Higgs boson turns out to be exactly what is predicted from the standard model, it will have zero internal angular momentum

(‘spin-0’), and a mass somewhere between 115 and 180 billion electron volts (GeV) in the energy units favoured by particle physicists. But such a match would be pretty boring, says John Ellis, a theoretical physicist at King’s College London. It would be much more fun, he thinks, if the LHC physicists didn’t find anything. “After all those years of speculating, we finally look under the appropriate lamp post, and it’s not there at all!” says Ellis. That would force the theorists back to the drawing board, “but there are various drawing boards to go back to”. There could be more complicated ways of generating mass, or something more unexpected. “That would be very exciting,” says Ellis.

Alternatively, he says, it is entirely possible that the LHC will turn up not one Higgs particle, but a whole family of them. That would be a sign of supersymmetry, a theory that predicts a zoo of as-yet unobserved ‘superparticles’, one to match each of the 25 particles in the standard model — the force-carrying bosons such as photons, gluons and the Higgs, and the fermions, such as quarks and electrons, which make up matter. These superpartners would be heavy — at least 600 GeV.

Supersymmetry appeals to physicists because it provides a unified mathematical description of bosons and fermions, which otherwise seem utterly unrelated. And the theory would greatly strengthen the case for a ‘grand unification’ of the strong, weak and electromagnetic forces, leaving only gravity unexplained.

In the standard model, the strong force’s interaction strength — expressed in terms of a constant analogous to electric charge — is very different from the strengths of the weak and electromagnetic forces. But if supersymmetry is assumed to be true, quantum corrections show all three strengths to be exactly equal — just as would be expected if the forces are actually one.

Supersymmetry would also solve some problems of other grand unified theories, such

as their predictions that the proton should be unstable. The presence of the superpartners in calculations tends to suppress proton decay, leading to a decay rate far below the limits currently set by experiment.

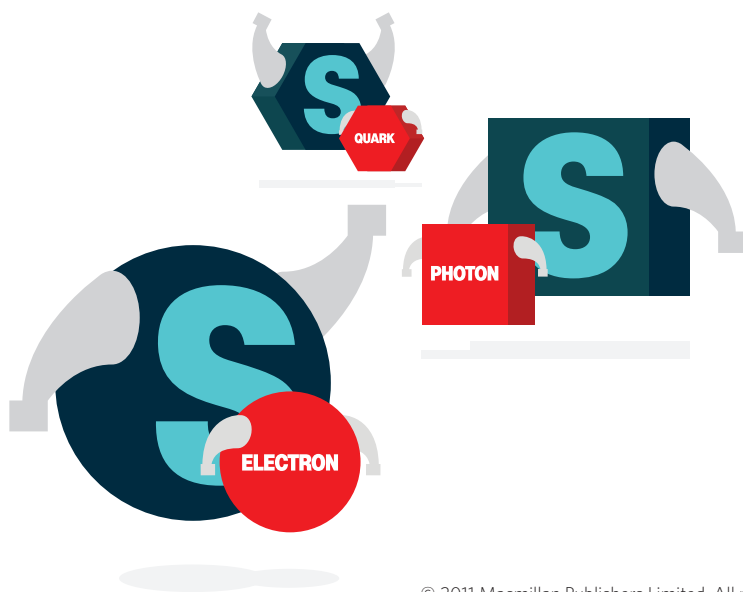
Finally, and perhaps most importantly from an observational standpoint, supersymmetry might very well provide an explanation for dark matter. This invisible cosmic haze behaves like a swarm of massive particles that interact very weakly with ordinary atoms, and has so far been detected only by its gravitational influence on visible stars and galaxies. No particle in the standard model has the right properties — but several of their predicted superpartners do. If one of them is indeed the dark-matter particle, it may soon be observed not only at the LHC, but also in one or more of the dark-matter detectors now operating around the world (see *Nature* doi:10.1038/news.2011.125; 2011).

But just because supersymmetry promises many wonderful solutions to current problems, that’s no guarantee it is true. “If supersymmetry is seen, that’s absolutely fabulous,” says Smolin. “But if not — well, that will be fabulous, too. It’s always better to know.”

EXTRA DIMENSIONS

Conversations with LHC physicists can become surreal — especially when they start talking with a straight face about finding that staple of science fiction, extra dimensions.

One big reason that they take this prospect seriously is that extra dimensions are predicted by string theory, by far the most popular attempt at unification beyond the standard model. String theory posits that the fundamental particles are actually vibrating threads of energy. Since it was developed in the late 1960s, string theory has remained a mental exercise, with no physical evidence to back it up. But it has proved remarkably compelling even so. It predicts the existence of forces that look a lot like the strong, weak and electromagnetic forces of the standard



model. It incorporates supersymmetry in a natural way. And it automatically includes gravity: string-theory equations show that closed loops of string would behave like gravitons, the particles postulated to carry the force of gravity. “It combines all the principles of physics we know,” says Nathan Seiberg, a string theorist at the Institute for Advanced Study in Princeton, New Jersey. “That’s huge. There’s no other suggestion that even comes close.” The extra dimensions arise because string theory has its most natural formulation in 11 dimensions, only 4 of which would be observable to us: the 3 dimensions of space and 1 of time. The missing dimensions are easy to explain, says Seiberg: the theory allows them to be so tightly rolled up that they’re invisible under ordinary circumstances.

The LHC could detect those extra dimensions if the particles generated by collisions have enough energy — and therefore short-enough quantum wavelengths — to start spiralling around those tightly curled dimensions. The energy of that spiralling would show up as mass, according to Einstein’s famous relativity theorem. So LHC physicists could detect whole families of higher-mass duplicates of the standard-model particles.

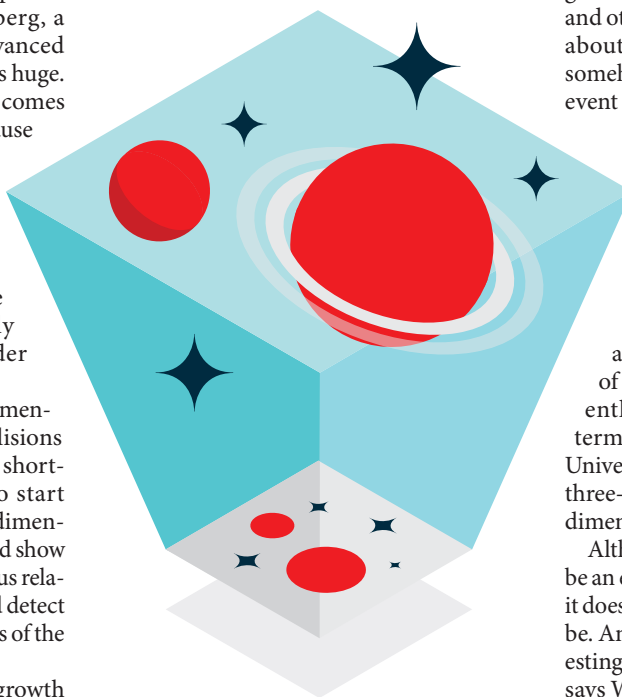
An alternative scenario, also an outgrowth of string theory, suggests that what we perceive as three-dimensional space is actually a kind of membrane floating in a higher-dimensional space. We never notice the extra directions because all the particles of the standard model are confined to the membrane. But LHC collisions might be energetic enough to let in a burst of gravitational energy from outside our membrane. The result would be a jet of collision products spraying off to one side of the collision point, but apparently with nothing to balance it on the other side — as if the jet had been hit by a bolt out of nowhere.

LHC physicists have calculated the experimental signatures of all these extra-dimensional phenomena and more, says Albert de Roeck, deputy spokesman of the Compact Muon Solenoid experiment, one of the LHC’s two big collision detectors. “But the moving target is to know the scale” of the extra dimensions, he says. If they are rolled up too tightly, on a scale smaller than 10^{-19} metres, then the energy required to probe them will be beyond the LHC’s reach.

That is a distinct possibility, which is why physicists consider it a long shot to observe extra dimensions at the LHC. “I’m not holding my breath,” says Michael Duff, a physicist at Imperial College London. “My bet is that if there are extra dimensions, they’re at the Planck scale” of 10^{-35} metres, where quantum mechanics and gravity are thought to unite in some still-unknown way.

Nonetheless, the pursuit of extra dimensions is considered well worth the effort.

Their existence would not prove string theory correct; it is perfectly possible to have extra dimensions without strings. But confirming one of its major predictions would considerably bolster the case for string theory.



HOLOGRAPHY

As nice as that would be, says Seiberg, string-theory practitioners are left with a gnawing sense that something is missing. “We know how to calculate a lot of things in string theory,” he says. “But we don’t have a conceptual basis for it — a set of fundamental principles from which everything follows.”

This quest for deeper principles is shared by physicists of all kinds, not just string theorists. One idea that has drawn a lot of attention follows from a startling theoretical discovery made by Stephen Hawking at the University of Cambridge, UK, and others in the 1970s: quantum effects in the space around a black hole cause it to emit radiation as if it were hot, even though black holes are supposed to swallow mass and energy, not spit it out. “That’s an amazing result,” says Carlo Rovelli, a physicist at the University of the Mediterranean in Marseilles, France. Somehow, the three apparently separate phenomena of gravity, quantum mechanics and thermodynamics — the science of heat — are intertwined, he says. “And we have still not understood why in a deep way.”

Efforts to understand this result have led theoretical physicists in some strange directions. According to standard thermodynamics, for example, any object’s temperature is related to its entropy: a quantity that measures the amount of information available to outside observers about the arrangement and motions of the object’s constituents. But

there is no way for outside observers to get any information from a black hole, because anything closer to the black hole than a surface known as the event horizon is cut off from the rest of the Universe. So if a black hole is governed by thermodynamics, as Hawking and others demonstrated, all the information about its three-dimensional interior must somehow be encoded on its two-dimensional event horizon.

Furthermore, after decades of analysis and generalization of this argument, many physicists now believe that it applies to any three-dimensional volume, from black holes to empty space: the volume’s entire information content can be encoded in its two-dimensional surface. Or to put it another way, the ultimate unified theory of everything should describe our apparently solid three-dimensional world in terms of a lower-dimensional reality. Our Universe would emerge from the theory like a three-dimensional optical image from a two-dimensional hologram.

Although this ‘holographic principle’ might be an element of some ultimate unified theory, it does not by itself say what that theory should be. And not every physicist buys it. “It’s interesting and provocative, but extremely vague,” says Wilczek. Mathematically, at least, a version of the holographic principle does apply in a string-theory model known as AdS/CFT duality, in which it has been studied extensively. And perhaps more importantly, it might be possible to test the idea.

Craig Hogan, director of the Fermilab Center for Particle Astrophysics in Batavia, Illinois, has suggested that if the holographic principle is true, quantum effects could produce a kind of ‘holographic noise’ in light beams. The effect would be minute, says Hogan, but might be detectable by the kind of ultraprecise laser interferometers already used in gravitational wave experiments.

He and his colleagues are already building hardware to test whether this effect will work as expected, says Hogan, and if it does, they hope to proceed next year with a full-scale experiment, projected to cost about US\$2 million over three years. “That’s cheap, as these things go,” says Hogan — and that is probably just as well. “These are tests of physics we don’t know, so by definition, it’s exploratory.”

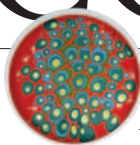
Among all the theories that could help to unify physics — from strings to holography and even more esoteric concepts — there is the possibility that many may turn out to be the same idea, viewed from different perspectives. But the only way to find out is to do the tests.

“The LHC gives us hope for a huge leap,” says Seiberg. “In a few years we should be a lot smarter.” ■ [SEE EDITORIAL P.265](#)

M. Mitchell Waldrop is an editor for *Nature* in Washington DC.

COMMENT

PHYSICS Maxwell's legacy — the promise of metamaterials **p.292**



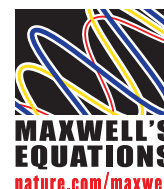
REVIEWS Size matters: eleven-page Spring Books special **p.294**

POLICY Intolerance of public concerns about science could be undemocratic **p.305**

OBITUARY Marshall Stoneham, leader on defects in solids, remembered **p.306**



As a young scientist, Maxwell wrote down the equations that laid the foundations of modern physics.



The laird of physics

James Clerk Maxwell's 1861 work on electromagnetism, which unified scientific fields, was driven as much by technology as by abstract theorizing, argues **Simon Schaffer**.

One hundred and fifty years ago this month, the 30-year-old James Clerk Maxwell published the first part of an extraordinary document: *On Physical Lines of Force*, a paper on the mathematical physics of electromagnetism. It is tempting to find in it much familiar furniture of the modern world: telecommunications, information technology and microelectronics. But the path from then to now is not as straightforward as it may seem. Maxwell's admirers sometimes treat his work as evidence that disinterested science eventually delivers huge economic and practical pay-offs. In the United Kingdom, for instance, 'Clerk Maxwell' has been proposed as the brand name for publicly funded regional centres for technology and innovation, with the ambitious goal of closing the gap between university and industry.

This linear model — the idea that cloistered theorizing can lead directly to new technology and profits — is often used by those arguing for greater investment in science. But a 2010 report from Britain's Royal Society titled *The Scientific Century* notes that "the

linear model of basic research through to innovation bears little relation to reality". It adds that the model's long-lasting appeal is partly explained by historical assumptions and familiar myths that are rarely challenged. Maxwell's magnificent work of the 1860s is an excellent example. Rather than a stately progression from abstract theory to solid application, it was the product of a web of markets, technologies, labs and calculators in the workshop of the world.

MYTHS AND MATHEMATICS

Why was Maxwell even working on electromagnetism? In the 1850s it was not common fodder for mathematicians educated at the University of Cambridge. The university's curriculum focused on the established sciences of celestial mechanics, wave optics and hydrodynamics. British universities did not even have teaching laboratories for physics. It is telling that the best-known British pioneer of electromagnetic studies was the Royal Institution professor Michael ▶

BETTMANN/CORBIS

► Faraday, a prodigious experimental chemist without formal education who knew little of higher mathematics.

Maxwell, unlike his hero Faraday, sometimes seemed a pastoral recluse. He often retreated to his 1,800-acre Scottish estate, where rural affairs absorbed much of his time — he was even described as a “north-country laird”. This cultivated self-image seemed to support the folklore of the brilliant scientist divorced from mundane trade.

Yet as a bright teenager in Galloway, in southwest Scotland, Maxwell had been unusually keen on turning classroom principles into practical devices, including homebuilt electric telegraphs. With his father’s earnest backing, Maxwell studied engines on show in the manufacturing towns and at the Great Exhibition of the Works of Industry of all Nations at the Crystal Palace in London in 1851. His ingenuity attracted the attention of William Thomson (later Lord Kelvin) — the two met in 1850. The young professor at the University of Glasgow encouraged Maxwell to manufacture magnetically sensitive crystals at home in Galloway.

After Maxwell completed his Cambridge mathematics training in early 1854, he understandably saw himself as “an electrical freshman”. But Faraday’s fascinating studies of electrical currents and magnetic poles drew his attention. So he asked Thomson, who had already been working on these puzzles, for guidance so that he and other graduates could at last “attack electricity”. The timing was perfect.

COMMUNICATION PROBLEMS

A project launched in the 1850s, aiming to tie the British Empire together with a global electric telegraph system, posed many of the problems that Maxwell and his scientific allies had to solve: the propagation of signals in electromagnetic networks, the reliability of measuring instruments, the analysis of force, resistance and current. It is, perhaps, significant that electromagnetic field theory was invented in the nation most concerned with submarine telegraphs. Elsewhere in Europe, overhead lines were commonplace, so problems caused by sea water in distorting and delaying signal transmission were not evident.

In 1853, Faraday witnessed the catastrophic problems of signal delay in cable lines laid between London and Manchester by a telegraph company. He saw these troubles as a chance to publicize his theory of electromagnetic induction. In a Royal Institution lecture a few months later, Faraday used his theory to show how sea water would act on submarine cables like the outer casing of a condenser, increasing the cable’s capacitance, slowing down induction and so delaying the signal even more.

Maxwell and Thomson closely studied Faraday’s lecture — and the warnings it held for submarine telegraphy. Within months Thomson

had a formula for the relationship between cable length and signal delay, of much interest to investors backing an ambitious scheme for a 5,000-kilometre undersea cable to North America. By 1856, Thomson had become a director of a major transatlantic cable company and, from his new Glasgow physics lab, a recognized authority on electromagnetic systems. In subsequent years, he earned many thousands of pounds from telegraph patents and consultancies, more than enough to buy a yacht, build a baronial mansion and endow his university with fellowships and equipment. Maxwell greatly admired Thomson for “bringing his obtrusive science to bear upon the engineers”.

Maxwell meanwhile, now a professor of natural philosophy in Aberdeen, boldly reworked Faraday’s model of electromagnetic induction. He reckoned that Faraday had destroyed fashionable continental models of electromagnetism, which posited instant actions across empty spaces between isolated particles. Maxwell began to see electromagnetism as dependent on some kind of fluid medium, a space-filling ether or field whose tension and stress stored energy and transmitted action at finite speeds. It took Maxwell four challenging

“British natural philosophers were trained to think in terms of the physics of pulleys, pumps, jelly or rubber.”

years from spring 1857 to make this model of electromagnetism match the best available data, first in Scotland, then at his new job as chair of natural philosophy at King’s College London. British natural philosophers, unlike their contemporaries in France and Germany, were trained to think in terms of the physics of pulleys, pumps, jelly or rubber, then extended to the Universe. This was

the approach Maxwell had learned in his studies of fluid mechanics at Cambridge. And so his paper of spring 1861 published in *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science* pictured the space of Faraday’s magnetic lines of force as a vast array of spinning gears separated by long strings of ball-bearings.

MECHANICAL EQUATIONS

Within this 1861 paper can be found, distributed through different sections of his argument, Maxwell’s first versions of the equations that characterize the electromagnetic field. Yet it was only much later that the equations were identified as fundamental building blocks of a new electromagnetic system. They summarized principles established by European experimenters in the previous decades. Electric flux through a closed surface is proportional to the electric charge it encloses; magnetic poles cannot exist in isolation; the electric current along a closed loop is related to the magnetic field round the loop; electromagnetic induction varies with time. To these principles Maxwell added the notion of a displacement current, generated even in apparently empty space by changing electrical forces.

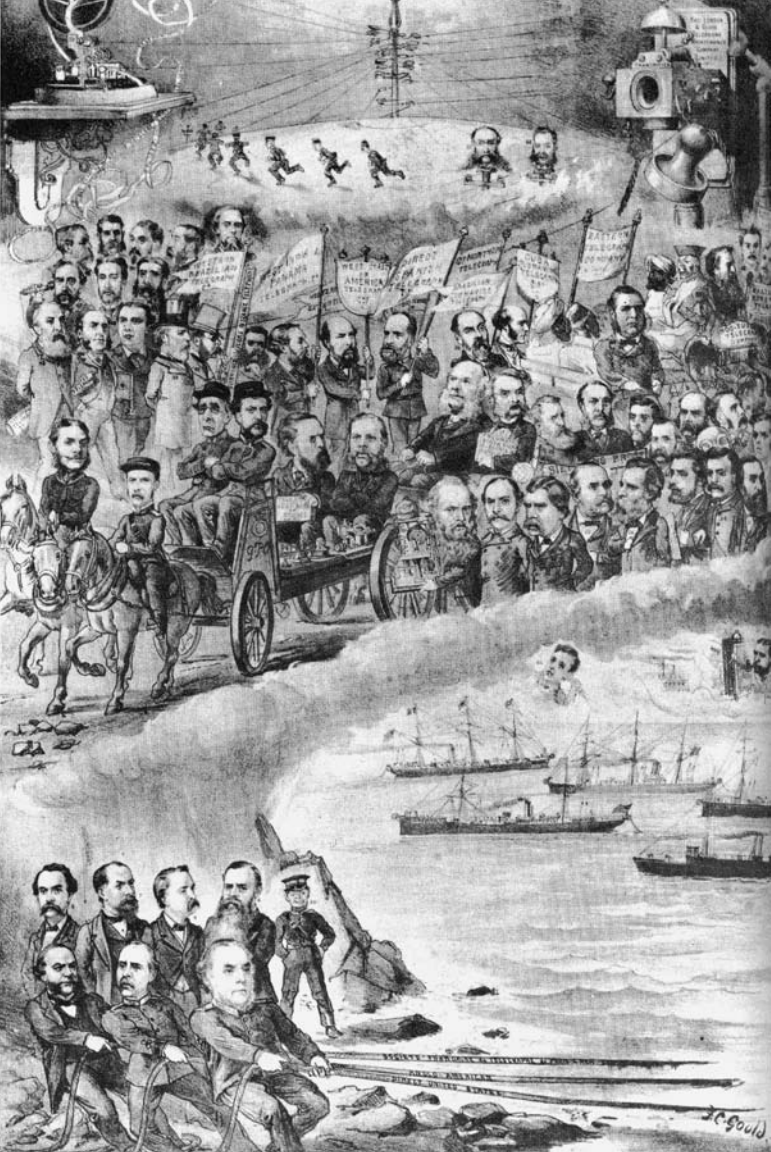
In his 1861 model, Maxwell showed that, when suitably adjusted, the workings of his complex machinery of gears and wheels matched those of electromagnetism. Magnetic energy was stored in the kinetic energy of the spinning gears, and the gears would spin and bulge, capturing the way changes in magnetic fields generated currents. With these achievements in print, Maxwell seemed entirely content when he went off to his country estate in summer 1861. In the next few months, however, his understanding of the significance of his own work changed significantly.

Over that summer, telegraph problems preoccupied British physicists. In the wake of the disastrous first attempt to lay a transatlantic cable, the British Association for the Advancement of Science set up a committee to determine standards of electrical resistance to make the cable system work. Within a year of the committee’s formation, Maxwell had joined enthusiastically, and electromagnetic measurements started to dominate Britain’s new physics teaching laboratories, including the one run by Thomson at Glasgow and, eventually, the Cavendish Laboratory led by Maxwell at Cambridge.

In Scotland in the summer of 1861, Maxwell thought harder about his model of fluid gears and wheels. Accurate technical measurements dictated some of the parameters of its structure. Maxwell realized



Sections of the Atlantic cables from 1858, 1865 and 1866.



An 1883 cartoon pictures physicists, engineers and bankers side by side.

that his fluid medium must transmit actions through transverse waves, just like rubber or jelly. He could calculate the waves' velocity from the density and rigidity of the medium — determined by practical laboratory measures of the electrostatic and electromagnetic forces. Back in London, he checked electromagnetic observations made in the late 1850s in Germany and plugged these values into his equations. To his apparent surprise, the ratio of the electrostatic and electromagnetic units differed by no more than about 1% from the value of light speed measured in Paris a decade earlier. "The coincidence is not merely numerical," he told Faraday in autumn 1861. It seemed that light was thus a vibration in the medium responsible for electromagnetism. With this insight, Maxwell launched the unification of the apparently disparate fields of light and electromagnetism and helped to change physics forever.

The reception of Maxwell's discoveries was also influenced by practical concerns with telegraphic signalling. In the amazingly fertile period before his early death in 1879, Maxwell strove to turn his mechanical model of the electromagnetic field into a generalized dynamic theory of electromagnetic action. Despite its central importance to physics and the modern understanding of all types of radiation, Maxwell's theory was slow to catch on.

One of its sternest critics was none other than Thomson. He never really forgave Maxwell for replacing the language of complex machinery with an abstract algebraic notation. He saw Maxwellian electromagnetism as "rather a backward step," if not a kind of "nihilism". One reason the two great scientists differed was because, from the first, they had different understandings of telegraph signalling. Maxwell held that Faraday's initial investigation of signal distortion needed a unified

treatment of induction, conduction and the surrounding medium. Thomson's models of telegraphy relied on separate analyses of conduction along the wires and induction through the insulator round them.

In 1884, five years after Maxwell's death, Thomson even publicly claimed that he'd known all about the electromagnetic theory of light long before, because of his analysis of signal propagation in submarine cables. It seemed Thomson was confusing light speed in the electromagnetic field, Maxwell's great topic, with signal speed in an underwater cable, Thomson's specialty. The Irish physicist George Francis Fitzgerald, one of Maxwell's young disciples, had to write in protest to *Nature* "to prevent what I find is a very common mistake". The link between telegraph engineering and the fate of electromagnetic theory lasted beyond Maxwell's lifetime.

In the preface to his 1873 masterpiece, *A Treatise on Electricity and Magnetism*, Maxwell explained his debt to telegraphy, arguing that it gave "a commercial value to accurate electrical measurements" and also allowed tests "on a scale which greatly transcends that of any ordinary laboratory", private or public. Indeed, telegraphy provided Maxwell with crucial support for his theory through measures of the electrostatic and electromagnetic units made in 1868–69.

Maxwell's decision to take over Cambridge's new experimental physics lab was in part driven by the need for such engineering resources. Happily, by the 1870s, Cambridge mathematics teachers were introducing students to submarine telegraphy in their new course on Maxwellian electromagnetism.

ORDINARY GROWTH

In sum, *On Physical Lines of Force* is an odd text to use as example of the unyielding purity of physical science. Maxwell's formulae did not appear in their most familiar form until almost 25 years after its publication. The four famous equations linking electromagnetic forces and fluxes owe their elegant and economical vector form to a brilliant London telegraphist, Oliver Heaviside. He published them in 1885 in *The Electrician*, a trade journal for electrical engineers and businessmen.

A veteran of the 1860s submarine cable schemes, Heaviside needed mathematical tools for managing energy transmission without distortion or loss. Unknown to the solitary telegraphist, a young German physicist, Heinrich Hertz, also established these simplified expressions for the field equations at the same time. Practical problems helped make these beautiful expressions, even in their most familiar, and apparently most abstract, mathematical form. For some time after 1885, they were not the Maxwell but the Hertz–Heaviside equations.

The French scientist Pierre Duhem saw this new physics as a sign of the rosbifs' industrial utilitarianism. "We thought we were entering the tranquil and neatly ordered abode of reason," Duhem complained of a Maxwellian textbook, "but we find ourselves in a factory." This was satirical, but in many ways it was spot on. There was an intimate and complex relation between the interests of scientific inquiry and commercial, industrial and technical enterprises. Victorians knew that well.

An 1883 cartoon (pictured), in a popular financial magazine, showed British physicists marching side by side with bankers, engineers, publishers and stockbrokers. Over the caption "telegraph and telephone mag(n)ates", the cartoon targeted the threat of Thomas Edison's new telephone to the cable system. In the last public lecture he gave, Maxwell showed a University of Cambridge audience this telephone. He had a performance of the song *Men of Harlech* transmitted from the Sedgwick Museum of Earth Sciences across town to the Senate House. And he taught the key lesson that the linkage of "these currents of the telephone which produce an audible effect" with precision electromagnetism was part of "the ordinary growth of scientific principles". Maxwell's triumphs of 150 years ago still have much to teach us about how this 'ordinary growth' works. ■

Simon Schaffer is in the Department of History and Philosophy of Science, University of Cambridge, Cambridge CB2 3RH, UK.
e-mail: sj16@hermes.cam.ac.uk



A new fish-eye lens
based on an idea of
James Clark Maxwell.

To invisibility and beyond

Combining Maxwell's equations with Einstein's general relativity promises perfect images and cloaking devices, explains **Ulf Leonhardt**.

Many everyday products of modern technology — such as mobile phones, television, computers and electric light — would seem almost magical to our ancestors. These all derive from James Clerk Maxwell's unification of the laws of electricity and magnetism 150 years ago. Scientists are still creating new wonders in the laboratory that exploit Maxwell's laws of electromagnetism. Some of these devices are made from extraordinary 'metamaterials' that can perform unusual tricks with light¹ — from optical cloaking to perfect imaging. Just last week, my colleagues and I announced evidence for 'perfect imaging'² using a device based on an idea³ of Maxwell's from 1854: a fitting tribute to a theorist who always thought in practical terms.

The field of metamaterials is barely ten years old. Early metamaterials relied on advances in nanotechnology to build tiny structures, such as metallic rings or wires, that are smaller than the wavelength of light. These nanostructures modify the electromagnetic properties of the metamaterial, sometimes creating optical effects that are not seen in nature. In 2006, for example, US scientists made the first prototype of an electromagnetic 'cloaking device'⁴. This makes a coin-sized object invisible to microwaves of a certain polarization and frequency. Modern metamaterials are also being used to make perfect lenses, that can image details finer than the wavelength of light.

What I find fascinating about metamaterials is how they connect my research area, optics, with Albert Einstein's theory of general relativity. The link is Maxwell's equations. It reminds me how much Einstein owed to Maxwell (a debt he always acknowledged), and also offers a way for the mathematical tools of general relativity to become practically useful in engineering.

In 1861, at the age of 30, Maxwell collected together all the laws of electricity and magnetism, and added one of his own. In doing so, he was the first to unify the concept of light with electricity and magnetism.

As he wrote in 1864: "We have strong reason to conclude that light itself — including radiant heat and other radiation, if any — is an electromagnetic disturbance in the form of waves." Without this insight, we would have no understanding of the electromagnetic spectrum, from radio waves and microwaves through the visible to X-rays and gamma rays.

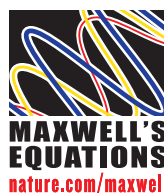
His discovery led to many subsequent advances in our understanding of light and matter. Einstein's theory of special relativity follows directly from Maxwell's equations in empty space — that much is obvious if you read Einstein's original papers. But it surprised me to learn how much Maxwell there is in Einstein's theory of general relativity (his theory of gravity) where, according to American physicist John Wheeler, "mass tells space how to curve and space tells mass how to move".

VIRTUAL SPACE

In fact the connections between Maxwell and Einstein are all around us — optical materials such as glass or water can also be said to curve space⁵. Looking through the front or top of an aquarium, for example, fish inside can appear at different positions and in different sizes, depending on your viewpoint. The glass and the water change your optical perception of space, but the fish still swim happily at their actual locations in physical space.

The aquarium creates a 'virtual' space for what our eyes see, different from ordinary physical space. This virtual space is the space experienced by light — not by the fish — and it has a curved geometry that corresponds exactly to that calculated using Einstein's theory of general relativity. So curved space doesn't just belong to cosmology — it is commonplace.

Rethinking optical materials as virtual spaces is an interesting idea, but is it useful in practice? It turns out to be crucial to the design of



YUNGUI MA/REF. 2

metamaterials and optical devices. Transformation optics¹, as the combination of general relativity and optics⁵ is now known, has inspired engineers, physicists and mathematicians to dream up many wonderful devices. Only a tiny fraction of these ideas will make it out of the laboratory, but some will do the seemingly impossible.

For example, cloaking with metamaterials is easy in theory⁶ (see 'Cloaking by transforming space'). In practice, it works only for specific wavelengths⁴. So an object may seem invisible to microwaves, but not to visible light. Scientists working in the visible spectrum have recently succeeded in creating devices that achieve partial cloaking, also known as 'carpet cloaking', whereby a three-dimensional object is made to look flat⁷. This feat of camouflage can be done with almost natural materials (silicon structures that look like tiny woodpiles) rather than complicated ring-and-coil metamaterials.

Carpet cloaking with silicon⁷ is related to my work, because I prefer to focus on materials with optical properties that are closer to those found in nature. These are not metamaterials, although they are engineered in ways that obey the rules of transformation optics. For such materials, a mathematical theorem — the Riemann mapping theorem — forbids cloaking transformations. It took me three years to find a way around the theorem⁸, but doing so allowed me to extend the framework of transformation optics to a broader range of wavelengths. However, we have not yet demonstrated real cloaking devices constructed from natural materials, and true invisibility for visible light remains impractical with existing materials and fabrication technology.

THE PERFECT IMAGE

In my opinion, the most promising potential of transformation optics lies in imaging. The idea of perfect imaging is what launched the field of metamaterials — and it takes us neatly back to Maxwell. Ordinary imaging devices such as microscopes suffer from a fundamental problem: they cannot image structures much smaller than half the wavelength of light, the diffraction limit. One cannot take a snapshot of atoms and molecules, because they are too small. In 2000, John Pendry introduced the concept of metamaterials⁹, showing that materials with negative refraction can, in theory, make a perfect lens that beats the diffraction limit. Materials with negative refraction can bend light in a direction that would not occur normally (all natural materials have positive refraction).

By 2006, Pendry's perfect lens could be understood by using the tools of transformation optics¹⁰. The lens appears to have folded space: a plane of physical space appears like a folded sheet of paper in virtual space. The electromagnetic waves in the folded regions are absolutely

identical, which explains why the image is a perfect copy of the original and no information is lost. But if light waves in two regions are identical, then light must instantly hop from one region of space to another, which, according to Einstein, cannot happen. Therefore, in reality, perfect imaging with negative refraction must be impossible for any useful device over practical distances. To get a negatively refracting perfect lens to work, imaging is limited to distances far smaller than the wavelength of light used.

An alternative approach can produce perfect imaging without using negatively refracting materials³; it is inspired by another idea of (who else?) Maxwell³. As a mathematics student at Trinity College at the University of Cambridge, UK, Maxwell dreamed up an optical device that reminded him of the eye of a fish. In Maxwell's 'fish-eye' lens, light travels in physical space as if it were confined to the surface of a virtual sphere. On this virtual sphere, light rays would go round in circles so that all light waves emitted from one point would meet again, perfectly, on the opposite side, just because of the symmetry of the sphere. Maxwell showed in 1854 that this fish-eye lens would give perfect resolution — that is, a point source appears as a point image. When Maxwell proposed his fish-eye lens he knew nothing about the wave nature of light — his electromagnetic discoveries were still five years away. And so for the past 150 years it was assumed that the wave nature of light would, in practice, restrict the resolution of a fish-eye lens to the diffraction limit.

In 2009 I argued, using transformation optics, that a fish-eye lens should in fact image waves with perfect resolution. As with Pendry's prediction⁹ of perfect imaging, this proposal created much controversy, in part because it sounds too good to be true and in part because it contradicts accepted wisdom. However, we have recently demonstrated perfect imaging for microwaves² using a two-dimensional version of the fish-eye lens (see image). We built our perfect lens using a metamaterial constructed from concentric bands of copper circuit board surrounded by a metallic mirror. Without a detector, the microwaves are reflected back and forth between source and image. But with a detector array in place (similar in principle to a digital camera) the fish eye can resolve two point sources that in ordinary imaging would appear blurred together. In principle, this route to perfect imaging should be achievable without using structured metamaterials. The next step is to demonstrate the same device for light, rather than microwaves.

What is truly remarkable about transformation optics is that by connecting Maxwell to Einstein, a theory as abstract as general relativity has actually become useful in engineering. Both men are known for their beautiful theories, but they were practical theoretical physicists: Maxwell performed experiments of his own, and Einstein always enjoyed making inventions and filing patents. Of course, in an ideal world run by wise politicians, we would not need to worry about justifying science as practical or fundamental. My mentor Stig Stenholm said that "the discovery of Maxwell's equations has already paid for all fundamental research for the following 500 years", because it laid the foundations of most of modern technology. We ought to have 350 years to go, no questions asked, thanks to Maxwell. ■

Ulf Leonhardt is a theoretical physicist at the University of St Andrews, Fife KY16 9SS, Scotland, UK.
e-mail: ulf@st-andrews.ac.uk

CLOAKING BY TRANSFORMING SPACE

Objects can be rendered invisible to certain wavelengths of light.

A cloaking device deforms the grid of space around its interior.

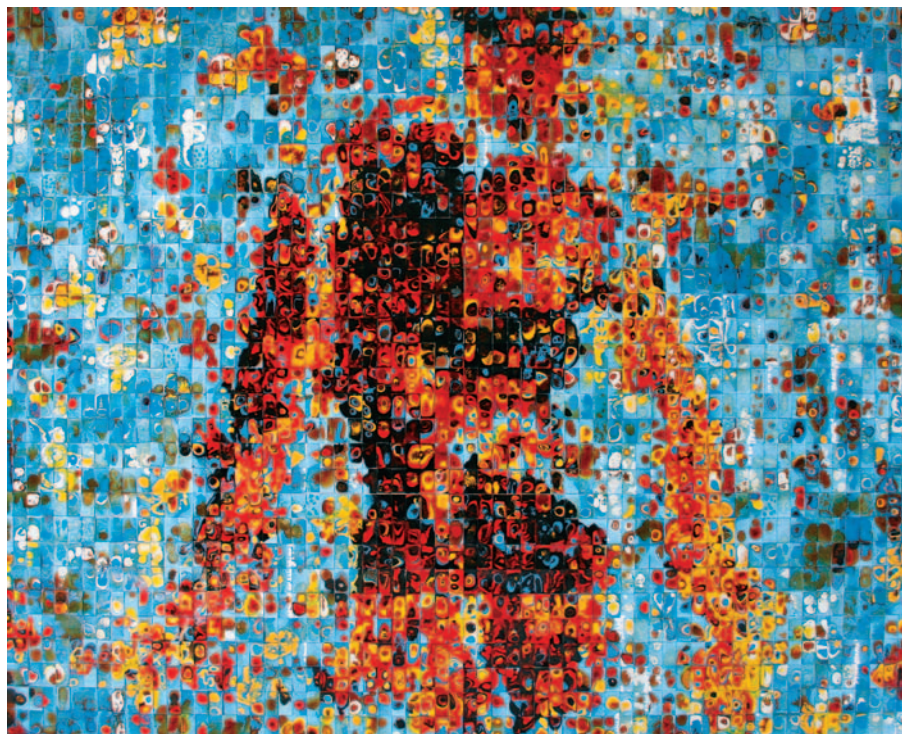
Light rays are bent as they travel through the device and exit it as if they have travelled through empty space.

Anything inside the device is hidden from an external viewer, and also doesn't affect the light rays carrying the image of the surrounding scenery, which creates the ultimate optical illusion: invisibility.

1. Service, R. F. & Cho, A. *Science* **330**, 1622 (2010).
2. Ma, Y. G., Sahebdivan, S., Ong, C. K., Tyc, T. & Leonhardt, U. *New J. Phys.* **13**, 033016 (2011).
3. Maxwell, J. C. *Camb. Dublin Math. J.* **8**, 188 (1854).
4. Schurig, D. *et al. Science* **314**, 977–980 (2006).
5. Leonhardt, U. & Philbin, T. G. *Geometry and Light: The Science of Invisibility* (Dover, 2010).
6. Pendry, J. B., Schurig, D. & Smith, D. R. *Science* **312**, 1780–1782 (2006).
7. Ergin, T., Stenger, N., Brenner, P., Pendry, J. B. & Wegener, M. *Science* **328**, 337–339 (2010).
8. Leonhardt, U. *Science* **312**, 1777–1780 (2006).
9. Pendry, J. B. *Phys. Rev. Lett.* **85**, 3966–3969 (2000).
10. Leonhardt, U. & Philbin, T. G. *New J. Phys.* **8**, 247 (2006).

SPRING BOOKS

COURTESY OF THE ARTISTS AND JAVAY, SACRAMENTO, CALIFORNIA



Size Matters: Figure 2 (2007) by Ian Harvey and Koo Kyung Sook.

EXHIBITION

Size Matters

San Jose Institute of Contemporary Art,
California.

Until 18 June 2011.

www.sjica.org

This Spring Books special issue displays a selection of works from *Size Matters*, an exhibition featuring ten North American artists who address ideas of size and scale. The works view the world from unusual perspectives, from Ian Harvey and Koo Kyung Sook's wall-sized enamel and shellac mosaics of human figures to the miniscule sculptures of Dalton Ghetti, carved from the graphite points of pencils. Expressed in a range of media, including photographs, paintings and video, the works comment on biological building blocks, knowledge, emotions and the environment.

BIOLOGY

A revolution in evolution

Manfred Milinski enjoys Martin Nowak's paean to the power of cooperation to shape animal and human societies.

Leading evolutionary theorist Martin Nowak sees cooperation as the master architect of evolution. He believes that next to mutation and selection, cooperation is the driving force at every level, from the primordial soup to cells, organisms, societies and even galaxies. Without cooperation, he says, our predecessors would still be RNA molecules. He sets out his groundbreaking ideas in *SuperCooperators*.

Co-authored with science journalist and editor of *New Scientist* Roger Highfield, *SuperCooperators* is part autobiography, part textbook, and reads like a best-selling novel. Nowak celebrates his oeuvre on the evolution of cooperation and challenges the mathematical basis for theories of kin selection and punishment. He is correct that this

part of evolutionary theory needs revisiting, but it is too soon to tell whether his bold ideas will hold up to empirical testing.

Game theory is central to Nowak's work and the book highlights five ways to work together for mutual benefit: direct reciprocity, indirect reciprocity, spatial games, group or multilevel selection and kin selection. Direct reciprocity is the tit-for-tat exchange of resources, which may be generous but is open to exploitation. Nowak believes that indirect reciprocity, where I help you and someone else helps me, is the most important mechanism driving human sociality. It enforces the power of reputation, gained

NATURE.COM
For another
book review on
cooperation, see:
go.nature.com/82iaph



**SuperCooperators:
Altruism, Evolution,
and Why We Need
Each Other to
Succeed**

MARTIN A. NOWAK WITH
ROGER HIGHFIELD
Free Press: 2011.
352 pp. \$27

by helping or refusing help, which is spread through gossip, thus selecting in evolutionary terms for sophisticated language. "Indirect reciprocity is the midwife of language and of our big, powerful brain," he says.

Cooperators can prevail through exchanges that are played out across and between networks and clusters of individuals, he explains. Multilevel or group selection follows among

communities that are small, numerous and isolated; mediated for example by tribal wars for resources. However, the migration of individuals between groups can undermine cooperation — egoists might take over pure altruist groups. *SuperCooperators* notes that there is plenty of evidence for group selection at the cellular level, such as strains of the bacterium *Pseudomonas fluorescens* that collectively produce a mat of polymer that allows the group to float on liquid surfaces.

More contentious is Nowak's approach to kin selection, or nepotism, in which individuals cooperate to ensure the success of genetic relatives in preference to strangers. Nowak set out his objections to this theory last year in a controversial *Nature* paper, co-authored with Corina Tarnita and Edward O. Wilson (*Nature* 466, 1057–1062; 2010). They question the theoretical basis of kin selection, or inclusive fitness theory: one of the cornerstones of the evolution of social behaviour.

Nowak and Highfield defend this view in *SuperCooperators*. After reviewing the history of evolutionary ideas about kin selection, including the lives of pioneering evolutionary theorists Bill Hamilton, George Price, John Maynard Smith and J. B. S. Haldane, Nowak criticizes key equations and calls them a recipe for disaster. He argues that the predictions of Hamilton's rule, which quantifies whether or not a gene for altruistic behaviour towards relatives will spread in a population, almost never hold. And he decries Price's fundamental equation, on which current inclusive fitness theory is based, as the mathematical equivalent of tautology.

In place of inclusive fitness theory, Nowak sketches a new model for the evolution of sociality, in which relatedness, he says, is a consequence rather than the cause of social behaviour. By assuming only one mutation — one that causes offspring to stay in the nest rather than leave — he claims to explain why progeny happen to be around to help their related mother. This model implies that offspring would help any unrelated elder in whose nest they were born, irrespective of a genetic link, and it does not explain why parents insist on caring for their own offspring rather than others. Here, in my view, relatedness is essential. Many experimental results support this, such as the sex ratios in colonies of different ant species.

In ant species in which the queen mates only once, for example, a preponderance of female reproductive offspring benefits the workers more than it does the queen: the non-reproductive workers support their mother to produce sisters, to which they are more closely related (75%) than is the queen (50%), thus more effectively perpetuating their genes than if they raised their own offspring. By contrast, in slave-maker ants, in which workers are stolen from other species and are therefore unrelated,



Size Matters: Detail from Figure 2 (2007) by Ian Harvey and Koo Kyung Sook.

the queen manipulates them to produce an equal sex ratio in her offspring for her own benefit. I anticipate that a better mathematical formulation of social evolution theory will be found that includes relatedness, is compatible with existing evidence and includes Hamilton's rule as a rule of thumb. Nowak himself states that "kin selection is a valid mechanism if properly formulated".

In another assault on established views, Nowak strongly disputes the effectiveness

NOWAK BELIEVES THAT COOPERATION HOLDS FOR 'ANY AND EVERY GAME IN THE COSMOS'.

of punishment as a method for promoting cooperation. Here he splits from his erstwhile colleague, game theorist Karl Sigmund, who accepts that the stick can be as useful as the carrot. Nowak, the theorist, describes how he performed experiments. In a version of the prisoner's dilemma game — in which two isolated players may choose to cooperate and both benefit, or one defects and receives a greater reward, being eventually punished by the other — he showed that those who do not punish gain most. No one has yet showed that punishers can gain from punishing, so it is not clear why punishing exists.

Nowak performed another experiment that, alas, failed to prove that reward rather than punishment promotes public cooperation. Clearly, the jury is still out on this question.

SuperCooperators is also Nowak's autobiography. After attending an all-boys school, he relates how he met his wife on his first female-dominated pharmacology course. And he recounts moments shared with his supporters: mountain climbing with chemist Peter Schuster; walking through the ancient forests of Austria's Rauriser Urwald with Karl Sigmund; playing soccer with theoretical ecologist Bob May; or dining on a Caribbean beachfront with Jeffrey Epstein, the Wall Street tycoon who funded Harvard University's Program for Evolutionary Dynamics, of which Nowak is director.

Nowak finishes with his concern for our planet, and of how Mahler's symphony *Das Lied von der Erde* (The Song of the Earth) carries a deep resonance for him. He worries about the climate game that everyone is now playing. "I believe that climate change will force us to enter a new chapter of cooperation," he writes, but his research does not provide a recipe.

A pleasure to read, *SuperCooperators* offers an explanation of the evolution of cooperation and shows where the experts disagree. Yet Nowak's faith in cooperation is so great that he believes his approach holds for "any and every game in the cosmos" — for all evolutionary processes on Earth, in our Galaxy and others, in "agglomerations of ancient stars that lurk in the faintest, farthest reaches." We will see. ■

Manfred Milinski is a director of the Max Planck Institute for Evolutionary Biology, Department of Evolutionary Ecology, Plön, Germany.
e-mail: milinski@evolbio.mpg.de

PHYSICS

Fundamental Feynman

An account of the physicist's work reminds **Leonard Mlodinow** of the gulf between theory and experiment.

In 1981, shortly after I arrived in the physics department at the California Institute of Technology in Pasadena, I heard a strong voice resonating down the corridor: "Hey Schwarz, how many dimensions are you in today?" The answer then was 10; it was once 26; it is now 11. Richard Feynman, who was teasing John Schwarz — one of the founders of string theory — didn't think much of any theory in which it wasn't four, for that is all we observe.

Quantum Man, by theoretical physicist Lawrence Krauss, focuses on the intimate connection that Feynman, like other physicists of his era, felt should exist between theories and experimental data. Writing to his third wife Gweneth from a gravitation and cosmology conference in Warsaw in 1962, he complained: "Because there are no experiments, this field is not an active one, so few of the best men are doing work in it. The result is that there are a host of dopes here (126) and it is not good for my blood pressure."

Today, gravitation and cosmology attract many of the best minds, who between them have produced so many unproven and competing theories of the multiverse that string theorist Brian Greene was able to write a long, popular account of them in *The Hidden Reality* (Allen Lane, 2011). Meanwhile, most scientists who study that cousin of cosmology, elementary particle theory, work in string theory despite its undetected dimensions and other apparent disconnects with reality, issues that current experiments cannot resolve.

By contrast, as Krauss recounts, when Feynman first presented his then-incomplete ideas on quantum electrodynamics (QED) describing the interaction between light and matter to the physics community in the late 1940s, he had "calculated almost every quantity one could calculate in QED" to ensure his results agreed with other methods and

experiment as far as was known. Much of the book concerns the intellectual journey that culminated in that work — a reformulation of quantum theory itself. It is a welcome addition to the shelf of Feynman biographies.

The story begins at Princeton University in New Jersey, with Feynman "in love" with the problem of the self-energy of the electron. Because like charges repel, each portion of a ball of negative charge exerts a repulsive force on every other portion. As a result, a ball of charge has a certain electric energy associated with it. The problem is, an electron is a point particle, but when you shrink the ball down to a single point, the repulsive energy becomes infinite.

Feynman thought that infinity had its roots in the way electromagnetic theory was formulated. This line of reasoning led him to recast the theory of electromagnetism in terms of an 'action principle', an exotic mathematical form that involves only the paths of charged particles over time, with no need to consider electric and magnetic fields. But the importance of what is now called Wheeler-Feynman electrodynamics (also named after US physicist

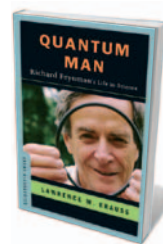


Size Matters: Large Petri Test 3 (2011) by Klari Reis.

John Wheeler) lies not in the theory itself, but in what it inspired Feynman to seek: a way to develop quantum mechanics around an action principle.

Feynman proposed a revolutionary new understanding of quantum reality. Imagine a particle that moves through some point A. According to classical physics, as it continues on its way, the particle will follow a definite path. Now consider another point, B. If B is positioned properly, the particle will eventually arrive there, but if B is located off the path, the particle won't. According to Feynman, the key difference in quantum theory is that the particle does not follow the classical path, or any single path. Rather, it samples every path connecting A and B, collecting a number called a phase for each one. Each of these, in concert, determines the probability that the particle will be detected at B. This novel approach, called the path integral or sum over paths, yields predictions equivalent to those of traditional quantum mechanics. Yet, as Feynman wrote, even if different theories are equivalent, "they are not psychologically identical when trying to move from that base into the unknown", meaning that they lead to different mental pictures, which can suggest different new ideas.

The unknown arena that Feynman moved into was the issue of how to fit quantum mechanics and special relativity into a single theory (we still don't know how to do this for general relativity, although string theory is a candidate). That synthesis, when applied to the electromagnetic force, is QED. In those days, QED, like string theory today, was a hard theory to make sense of. Quantum-physics pioneer Wolfgang Pauli wrote: "The risk is very great that the entire affair loses touch with physics and degenerates into pure mathematics." But in this case, Krauss points out, there were plenty of experimental data to guide and inspire Feynman, and after years of work and thousands of pages of calculations, he built a consistent and infinity-free theory



Quantum Man: Richard Feynman's Life in Science

LAWRENCE M. KRAUSS
W. W. Norton: 2011.
368 pp. \$24.95, £19.99

NEW IN
PAPERBACK

Highlights of this
season's releases



Perfect Rigor: A Genius and the Mathematical Breakthrough of the Century

Masha Gessen (Icon Books, 2011; £14.99)

In 2002, reclusive mathematician Grigori Perelman solved the Poincaré conjecture, one of the world's greatest intellectual puzzles. Shunning all publicity, he refused to accept the prestigious Fields Medal for his achievement and vanished from the public gaze. Journalist Masha Gessen attempts to discover more about him by travelling to Russia to interview Perelman's colleagues and teachers and discussing his behaviour with psychologists.

COURTESY OF THE ARTIST

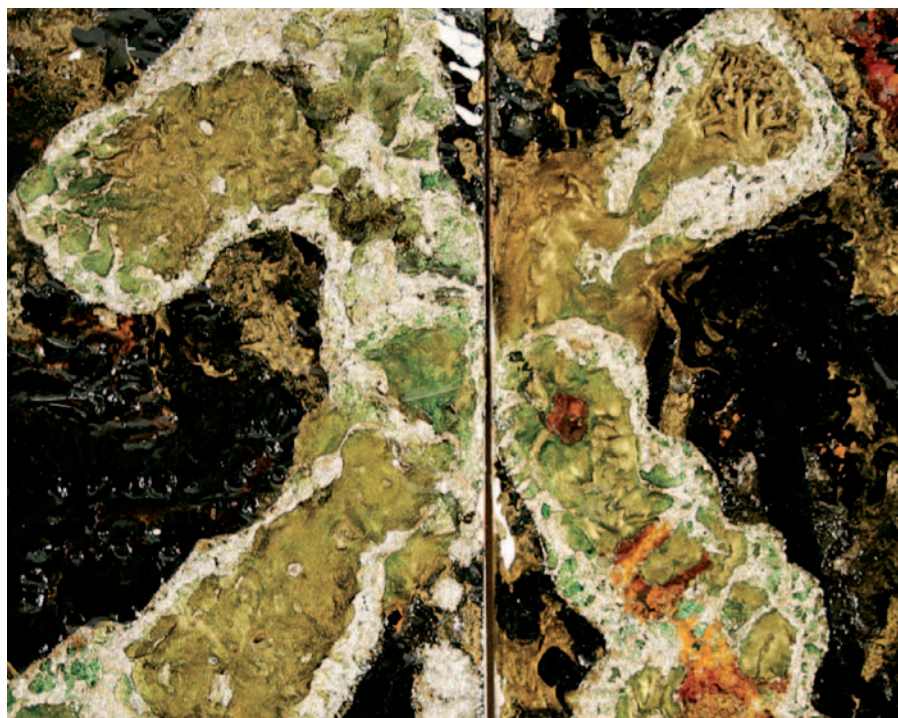
of QED. The result was so powerful that, according to US-based physicist Freeman Dyson, Feynman could do a calculation that once took several months “in half an hour” on the blackboard.

When he won his Nobel prize in 1965, Feynman felt his methods were merely useful, not profound. Today, his approach is considered a more fundamental way to look at quantum theory than the formulations of its founders, such as Niels Bohr, Werner Heisenberg and Erwin Schrödinger. It is the basis of how physicists think about particles interacting, exchanging carriers of force, fluctuating in and out of existence. It is also, ironically, the basic tool of both string theory and quantum cosmology.

Krauss does a good job of imparting Feynman's fascination with all physical phenomena, and goes on to describe Feynman's later groundbreaking work in other fields — on the weak interactions, the theory of liquid helium and his parton model, which provided evidence for the existence of quarks. Intertwined with the physics are snippets of Feynman's personal life, including his habit of working on physics in a strip club, and his undying love for his first wife Arline Greenbaum, who died from tuberculosis in 1945, just a few years after they were married.

As Krauss acknowledges, the book contains little that is new. Jagdish Mehra's *The Beat of a Different Drum* (Clarendon Press, 1994) is a far more detailed account of Feynman's science, equations and all; and James Gleick's best-seller *Genius* (Little, Brown, 1992) covers Feynman's personal life in greater depth. Still, I found the account of Feynman's hard work, passion and discoveries inspirational, and, for a physicist at least, good bedtime reading. For those without a strong physics background, however, the prose can be tough going, especially the more technical passages. Personally, I love being talked to that way. ■

Leonard Mlodinow is author of *The Grand Design* (co-authored with Stephen Hawking), *The Drunkard's Walk* and *Feynman's Rainbow*.
e-mail: len@caltech.edu



Size Matters: Detail from *Figure 6* (2008) by Ian Harvey and Koo Kyung Sook.

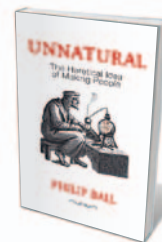
BIOTECHNOLOGY

Making people

Today's wariness of reproductive technologies stems from myths, legends and Hollywood, finds **Chris Mason**.

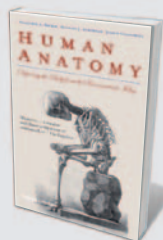
Stem cells, cloning, regeneration and life extension are frequently in the news. When they are, the media often resort to sensationalist clichés — invoking Frankenstein to conjure up a stereotypical mad scientist ‘playing God’ by creating out-of-control monsters. Whereas the creation of non-human artificial life, such as Craig Venter's engineered microbes, gets a mixed press, the making of humans is invariably controversial. Clearly, human life has a special moral status.

In *Unnatural*, science writer Philip Ball explores the history of our fascination with — and fear of — creating artificial people, from ancient folklore to today. Tracing a clear path



Unnatural: The Heretical Idea of Making People
PHILIP BALL
Bodley Head: 2011.
384 pp. £20

from medieval alchemists' homunculi to routine assisted conception is a feat. Through his impeccable research, Ball successfully argues that the tenacious myths of the past that surround the making of people or ‘anthropoeia’ (his coinage) affect life-science research today.



Human Anatomy: Depicting the Body from the Renaissance to Today

Benjamin A. Rifkin, Michael J. Ackerman and Judith Folkenberg (Thames & Hudson, 2011; £19.95)
This beautifully produced book presents more than 500 years of anatomical illustration. It charts how our knowledge about the body has changed along with our interpretation of what we see within it.



Wetware: A Computer in Every Living Cell

Dennis Bray (Yale Univ. Press, 2011; \$18)
By treating a single-celled organism as a computational system, biologist Dennis Bray explains how it balances internal chemistry, responds to light and hunts prey — all without a nervous system. He sees cells as unique molecular circuits that perform logical operations.

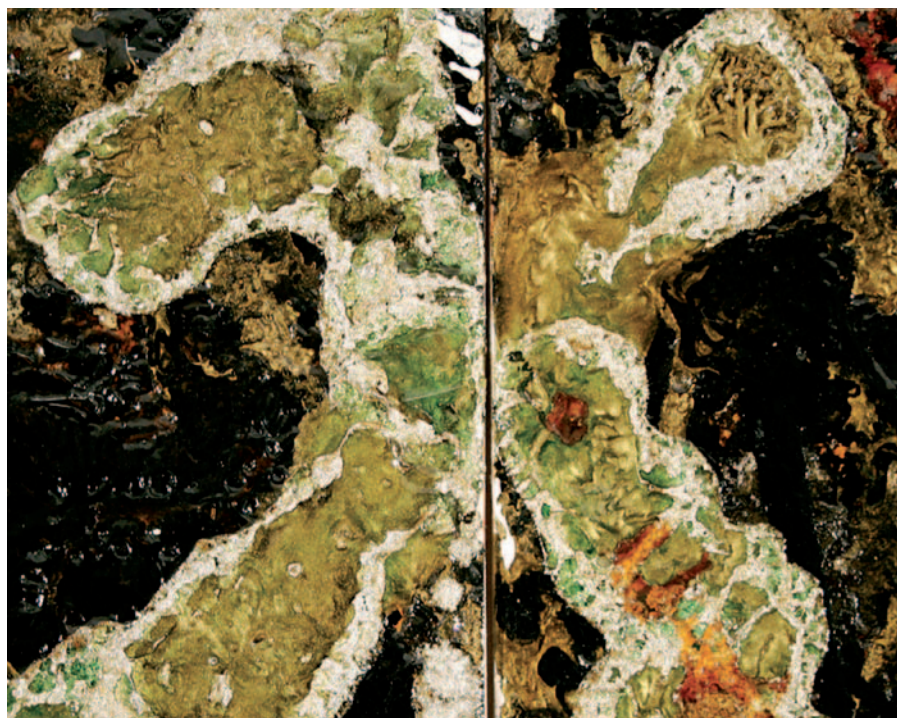
of QED. The result was so powerful that, according to US-based physicist Freeman Dyson, Feynman could do a calculation that once took several months “in half an hour” on the blackboard.

When he won his Nobel prize in 1965, Feynman felt his methods were merely useful, not profound. Today, his approach is considered a more fundamental way to look at quantum theory than the formulations of its founders, such as Niels Bohr, Werner Heisenberg and Erwin Schrödinger. It is the basis of how physicists think about particles interacting, exchanging carriers of force, fluctuating in and out of existence. It is also, ironically, the basic tool of both string theory and quantum cosmology.

Krauss does a good job of imparting Feynman’s fascination with all physical phenomena, and goes on to describe Feynman’s later groundbreaking work in other fields — on the weak interactions, the theory of liquid helium and his parton model, which provided evidence for the existence of quarks. Intertwined with the physics are snippets of Feynman’s personal life, including his habit of working on physics in a strip club, and his undying love for his first wife Arline Greenbaum, who died from tuberculosis in 1945, just a few years after they were married.

As Krauss acknowledges, the book contains little that is new. Jagdish Mehra’s *The Beat of a Different Drum* (Clarendon Press, 1994) is a far more detailed account of Feynman’s science, equations and all; and James Gleick’s best-seller *Genius* (Little, Brown, 1992) covers Feynman’s personal life in greater depth. Still, I found the account of Feynman’s hard work, passion and discoveries inspirational, and, for a physicist at least, good bedtime reading. For those without a strong physics background, however, the prose can be tough going, especially the more technical passages. Personally, I love being talked to that way. ■

Leonard Mlodinow is author of *The Grand Design* (co-authored with Stephen Hawking), *The Drunkard’s Walk* and *Feynman’s Rainbow*.
e-mail: len@caltech.edu



Size Matters: Detail from *Figure 6* (2008) by Ian Harvey and Koo Kyung Sook.

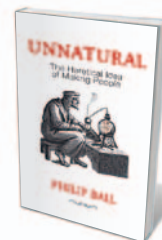
BIOTECHNOLOGY

Making people

Today’s wariness of reproductive technologies stems from myths, legends and Hollywood, finds **Chris Mason**.

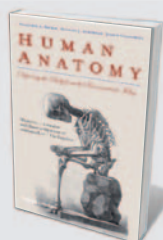
Stem cells, cloning, regeneration and life extension are frequently in the news. When they are, the media often resort to sensationalist clichés — invoking Frankenstein to conjure up a stereotypical mad scientist ‘playing God’ by creating out-of-control monsters. Whereas the creation of non-human artificial life, such as Craig Venter’s engineered microbes, gets a mixed press, the making of humans is invariably controversial. Clearly, human life has a special moral status.

In *Unnatural*, science writer Philip Ball explores the history of our fascination with — and fear of — creating artificial people, from ancient folklore to today. Tracing a clear path



Unnatural: The Heretical Idea of Making People
PHILIP BALL
Bodley Head: 2011.
384 pp. £20

from medieval alchemists’ homunculi to routine assisted conception is a feat. Through his impeccable research, Ball successfully argues that the tenacious myths of the past that surround the making of people or ‘anthropoeia’ (his coinage) affect life-science research today.



Human Anatomy: Depicting the Body from the Renaissance to Today

Benjamin A. Rifkin, Michael J. Ackerman and Judith Folkenberg (Thames & Hudson, 2011; £19.95)
This beautifully produced book presents more than 500 years of anatomical illustration. It charts how our knowledge about the body has changed along with our interpretation of what we see within it.



Wetware: A Computer in Every Living Cell

Dennis Bray (Yale Univ. Press, 2011; \$18)
By treating a single-celled organism as a computational system, biologist Dennis Bray explains how it balances internal chemistry, responds to light and hunts prey — all without a nervous system. He sees cells as unique molecular circuits that perform logical operations.

Ball traces the concept that nature is good and *techné* is bad back to Aesop's and Ovid's Prometheus, maker of humanity from earth and water, and provider of technology to man. After Prometheus came recipes for making miniature humans called homunculi. Starting in the Middle Ages, initially as a cure for childlessness, the art of homunculi-making evolved into a debate over whether the miniscule men had a soul. Johann Wolfgang von Goethe's nineteenth-century poetic play *Faust* raises this spectre. Deploying the biological equivalent of alchemy, Faust's former assistant, Wagner, creates his homunculus: a tiny super-being with magical powers who is trapped in a glass vessel, doomed to remain captive without the capacity to become a proper man. In 1818, Mary Shelley published *Frankenstein*, appropriately subtitled 'The Modern Prometheus', in which her eponymous scientist unintentionally constructs a monster, by unexplained means, from human parts. There are also golems — the animated beings of Jewish folklore, made from clay and brought to life by religious magic for the purpose of imitating God's creation.

Ball distills out of all this a set of universal myths surrounding anthropoeia that are deeply ingrained in society, resulting in the widely held view that artificial people-making is unnatural and deeply wrong — heretical, as in the book's subtitle. His thesis is that humans fear that uncovering forbidden knowledge will result in either divine or other retribution. Prometheus, Faust and Frankenstein all pay a heavy price for their transgressions into anthropoeia. Even today, Ball points out, societal and cultural debate is pervaded by the belief that technology is intrinsically perverting and thus carries certain penalty. Views that human cloning will be used for social engineering, eradicating one gender or resurrecting undesirable figures from the past, for example, all reflect age-old fears about the consequences of meddling in the 'unnatural'. Ball warns that, as there is no

global ban on human reproductive cloning, there is a strong chance that it will happen. It is thus likely to become a de facto reality without the well-informed debate it deserves.

As scientific knowledge accumulates and makes some acts of anthropoeia more and more plausible, the challenge for the public will be to separate fact from fiction. For example, Ball ends his literary tour with Aldous Huxley's novel *Brave New World*. In 1931, the book's *in vitro* production of embryos in the Central London Hatchery and Conditioning Centre was pure conjecture by Huxley, based on the scientific forecasts of his day. Today, *in vitro* fertilization (IVF) is mainstream medicine — more than four million babies have been born using this technique. But the technology

still has its critics, including within the Vatican. On the awarding of the 2010 Nobel Prize in Physiology or Medicine to IVF pioneer Robert Edwards, Ignacio Carrasco de Paula, head of the Pontifical Academy for Life, stated that the award was "completely out of order", as without IVF there would be no market for human eggs "and there would not be a large number of freezers filled with embryos in the world". For some, such words conjure up images of unscrupulous profiteering and factory-like storage of human lives, generating fears that human procreation will be reduced to mere money and industrial bioprocessing.

Huxley was more futuristic in including humans conceived and grown entirely outside the body. As Ball explains, the artificial womb remains fiction, albeit moving slowly towards fact. Its leading exponent, Hung-Ching Liu, at the Center for Reproductive Medicine and Infertility in New York, has grown human uterus lining (endometrium) and thinks it will eventually be possible for fetuses to be grown outside a woman's body. Progress has been

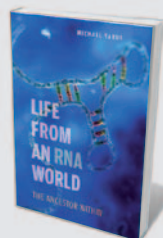
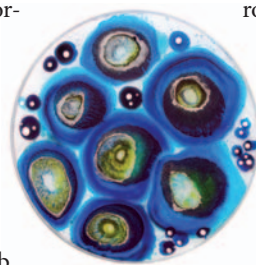
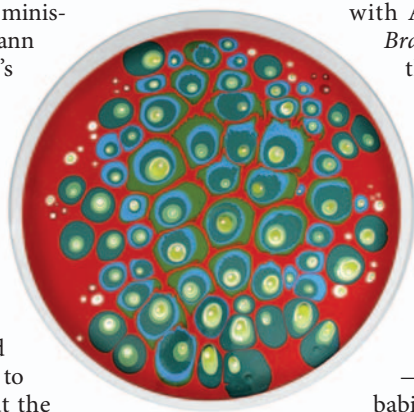
made in other species, including mouse embryos gestated to almost full term in 'bubbles' of endometrial tissue and premature goats kept alive by artificial placentas.

Meanwhile, headlines about three-parent human embryos and mice with two fathers continue to fuel science fiction. Back in 1978, the film *Boys from Brazil* imagined Nazi physician Josef Mengele attempting to resurrect Adolf Hitler by reproductive cloning. Some 20 years later, in *The Matrix* (1999), countless humans are bred and kept in pods so that their body heat and electrical activity can be harvested as energy for the machines that have taken over the world. More recent films, such as *The Island* (2005) and *Splice* (2009), have further built on the science fiction of reproductive science.

The challenge for innovative biological research is that, until it translates into real benefits, it is often viewed with mistrust and worse-case scenario imagery. In reality, once products and services are released into society, they are adopted by a few enthusiasts and then, if successful, by the wider community. In the 1970s, for example, anxieties were rife about the unfounded threat that IVF posed to human welfare and dignity, let alone whether a test-tube baby could ever be wholly human. Yet the first IVF baby, Louise Brown, was just like everyone else, so IVF became socially acceptable. We cannot predict whether human cloning will proceed in the same manner, so the past is our only pointer.

For scientists, clinicians and biotechnology business people, understanding deep-rooted ideas, however irrational, is vital for successful dialogue with the public. The fiasco of genetically modified (GM) crops came about because of the failure to predict that the media would label GM products as 'Frankenfood' — together with the moral judgement it would infer. Today, stem cells and cloning are under the media spotlight. *Unnatural* is therefore a must-read for all stakeholders of these advanced technologies. ■

Chris Mason is a professor at the Advanced Centre for Biochemical Engineering, University College London, UK.
e-mail: chris.mason@ucl.ac.uk



Life from an RNA World: The Ancestor Within

Michael Yarus (Harvard Univ. Press, 2011; \$17.95)
Many biologists think that today's DNA-based life forms evolved from RNA molecules. Biochemist Michael Yarus marshals arguments in support of that theory in his book, which explores the principles of Darwinian evolution, the tree of life and the diverse abilities of RNA.



Biology is Technology

Robert H. Carlson (Harvard Univ. Press, 2011; \$21.95)
Robert Carlson explains how to build synthetic biological systems from basic components and the technology used to manipulate them. "An informative view of the future prospects for biotechnology and its regulation," wrote reviewer Michael Goldman (*Nature* **464**, 1129–1130; 2010).

ANATOMY

How to get ahead

A thorough review of how the human head evolved shows how hominins outpaced apes, finds **Henry Gee**.

Almost everything we know of our world comes through one structure — the head. Our organs of sight, hearing, smell, taste and balance huddle in this crowded tenement, which comprises just 8% of our body's mass. Meanwhile, down in the basement, air flows in and out; food is swallowed and processed; and sounds emerge — from belches to Beethoven. All this activity is coordinated by the brain, the head's largest and most mysterious organ.

Given this foam of business, it is amazing that we get any peace. That we do is testament to the integration of the head — despite sharing such a small space, all its tenants get on famously. This integration, says Daniel Lieberman in *The Evolution of the Human Head*, his thorough review of the head's anatomy and development, is key to the powerful ability of the human head to evolve. A small amount of tinkering in one part can lead to a more comprehensive reorganization, as the various sections of the head mould themselves to new circumstances.

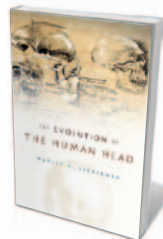
The head divides into three parts: the brain and the cranial vault that houses it; the cranial base on which the brain sits; and the face, hafted to the front of both. Each influences the others, explains Lieberman, a professor of human evolutionary biology at Harvard University in Cambridge, Massachusetts. Together the three parts determine the look of modern humans compared with other apes. A larger brain needs a more flexed cranial base to support it, thus shortening the face, which is retracted under the braincase. So our heads appear less extended than those of our closest animal relatives.

A shorter pharyngeal cavity relative to the neck gives us a system of tubes that produces more intelligible speech sounds than a longer

one could. Our short, round tongue decouples the epiglottis from the soft palate, giving more access for odorants to nasal epithelia. A shorter face means a smaller jaw, all the better for chewing cooked foods, thus getting more energy per morsel, necessary for maintaining a large brain. This positive feedback loop contributes to the continuing success of *Homo sapiens*.

Lieberman does well to steer a course away from the blurry and disputed details of hominin lineages. Instead he focuses on the basics of how the head develops. He sets out the elegant certainties of biomechanics: how the strains of chewing influence bone deposition and resorption; how the mechanics of jaw shape interact with the cranium and the associated tendons and musculature to integrate the head's components; how the shape of the nose promotes turbulent airflow and thus efficient detection of odorants, and moderation of heat loss and water resorption; and how the dimensions of the ear canal attenuate some frequencies while amplifying others.

Lieberman's thoroughness especially enriches the final quarter of the book — a tour of human evolution in terms of the detailed changes wrought on head anatomy over the past few million years. Climate change was to blame: the drying, cooling climate forced forest apes to either



The Evolution of the Human Head
DANIEL E. LIEBERMAN
Belknap Press: 2011.
768 pp, \$39.95

retreat farther into the trees or adopt more marginal lifestyles.

A bipedal life, however, forced certain compromises. The lack of lengthy canines in early hominins might be less to do with sexual selection than the imposition of a shorter face (with less room for long teeth) on a head seeking balance atop an upright carriage. Some early hominins became better at subsistence on the marginal fare of roots and seeds, evolving immense teeth and powerful, crushing jaws. Others, the descendants of which became *Homo*, found that although nutritious fruits were scarce on the savannah, nutritious ungulates were common — but they needed to be caught.

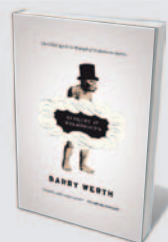
The hunting imperative led to a suite of features suggesting that humans, uniquely among primates, became excellent long-distance runners. No chimpanzee, for example, could attempt a marathon across the Sahara Desert. Yet human athletes do this, and survive. Running imposes many constraints and demands on the head of a biped, connected with stabilization and protection from shock, which are seen in the anatomy of humans but not in apes. One such feature, the nuchal ligament that links the back of the skull to the neck, finds parallels in other chasers-of-prey such as dogs.

Most pursuit predators and scavengers work at dusk, dawn or at night. Even before the evolution of Englishmen, early humans found that the midday Sun offered a vacant niche. Because humans can maintain a pace that exceeds the walking speed of many quadrupeds, even in the heat of the day, the quarry is forced to stand its ground or run and risk collapsing from hyperthermia. Bushmen adopt such a hunting strategy to this day. Its use by early *Homo* could have promoted the many adaptations seen for long-distance running, as well as the thoroughgoing changes in the structure of the head and face that set humans apart from all other apes.

By rooting his study in the basics of tissue mechanics and functional morphology, Lieberman does the spadework to which all such studies aspire but few achieve — and makes that task seem elegant and effortless. ■

Henry Gee is a senior editor at Nature.
e-mail: h.gee@nature.com

NATURE.COM
For a review of Neil Shubin's *Your Inner Fish*:
go.nature.com/yb2sk8



Banquet at Delmonico's: The Gilded Age and the Triumph of Evolution in America

Barry Werth (Univ. Chicago Press, 2011; \$19)
Philosopher Herbert Spencer took Charles Darwin's ideas to the United States in 1882. Barry Werth focuses on the influential diners at a banquet held in Spencer's honour, describing how they used evolutionary ideas in an attempt to improve society.



Mothers and Others: The Evolutionary Origin of Mutual Understanding

Sarah Blaffer Hrdy (Belknap Press, 2011; \$19.95)
In her provocative book, anthropologist Sarah Blaffer Hrdy argues that because human infants are too expensive to be raised by mothers alone, both parents must invest heavily in social skills to bargain with other group members for resources.



TECHNOLOGY

The medium is the message

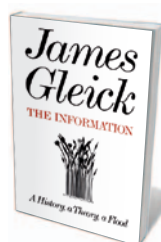
Thomas J. Misa enjoys a history of communication tools, from talking drums to Twitter.

Information is the paradigm of our time. Economies that once manufactured goods now create value by processing information. Global flows of money, ideas and news determine which countries engage with global society and which are left silently on the sidelines. A skein of pervasive mobile computing keeps us connected — instantly, continuously, incessantly. James Gleick's latest book, *The Information*, examines the genesis of the information society and the roots and consequences of information theory.

Gleick is no stranger to demanding scientific topics. His blockbuster *Chaos* (Penguin, 1987) popularized Edward Lorenz's mathematics of complexity. He is also the biographer of physicists Richard Feynman and Isaac Newton. In *The Information*, he highlights the great surge of classifying and calculating often labelled as the industrial and scientific revolutions, and he profiles leading theorists, notably US mathematician Claude Shannon.

Gleick acknowledges that the concept of information and its impacts are difficult to grasp, yet explains our fascination with seeing information as the driver of just about everything.

Rather than telegraphs or telephones, Gleick begins with 'talking' African drums. Because African languages had hundreds of sounds, it seemed impossible to European observers that complex messages could be conveyed using drums that made only two sounds, pitched high and low. Yet for centuries, almost all African people could understand the messages that were broadcast by skilled drummers.

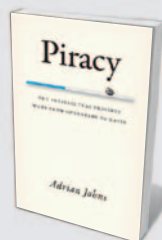


The Information: A History, A Theory, A Flood
JAMES GLEICK
Pantheon/Fourth Estate: 2011. 544 pp.
\$29.95/£25

After decades of European puzzlement, John Carrington's 1949 book *The Talking Drums of Africa* revealed all. There was no telegraph-like Morse code within drumming. African languages relied only partly on unitary sounds or 'phonemes' and more fundamentally on their intonation. Simply altering their tones could transform the phonemes for 'he watched the riverbank' into 'he boiled his mother-in-law'. With drum tones expressing the rising and falling pitches of African speech, drummers could accurately convey a complex message. And anyone whose ear was attuned could understand it. Using this accessible analogy, Gleick deftly introduces the concepts of information channels, intentional redundancy of messages and the importance of error correction.

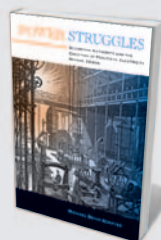
Gleick's more technical treatment of Shannon and information theory is a recurring thread of the crucial middle chapters. He serves up enlightening side views

COURTESY OF THE ARTIST; PHOTO: S. HOWARD



Piracy: The Intellectual Property Wars from Gutenberg to Gates

Adrian Johns (Univ. Chicago Press, 2011; \$22.50)
Intellectual piracy, as historian Adrian Johns explains, is nothing new. From the invention of the printing press to modern file-sharing, Johns explores the wars that have arisen over intellectual property rights. (See Steven Shapin's review: *Nature* **466**, 563; 2010.)



Power Struggles: Scientific Authority and the Creation of Practical Electricity Before Edison

Michael Brian Schiffer (MIT Press, 2011; \$19)
Behavioural archaeologist Michael Schiffer investigates electricity technologies before Thomas Edison's success. He shows why some made an impact while others failed, and the role of scientific authority in determining their fate.



Size Matters: Alphabet (2005) by Dalton Ghetti.

to Babylonian mathematics, the *Oxford English Dictionary*, Charles Babbage's mechanical computers, telegraph codes, the 'completeness' of formal mathematics, wartime cryptography and especially the telephone system, which provided a focus for Shannon's work. He also gives apt summaries of Walter Ong, Marshall McLuhan and other commentators on the information age.

The narrative of Shannon's place in information theory is well known to historians. Shannon, a distant relative of Thomas Edison, grew up in rural northern Michigan, studied electrical engineering and mathematics at the University of Michigan, then went to the Massachusetts Institute of Technology (MIT) in Cambridge. Here, after operating Vannevar Bush's massive mechanical differential analyser and working at Bell Telephone Laboratories during the summer break, he started work on his master's thesis: 'A Symbolic Analysis of Relay and Switching Circuits'. It was accepted by MIT in 1937 and published a year later. Psychologist Howard Gardner called it "possibly the most important, and also the most famous, master's thesis of the century". Telephone switching systems at the time were composed of thousands of electromechanical relays; within two decades or so, they

had evolved into electronic computers.

Shannon was the link between algebra and switches. He saw that the on-off states of telephone relays resembled the algebra originally conceived by George Boole in the 1850s, with its now-familiar notation of ones and zeroes and 'and', 'or' and 'not' operators, and that immense systems of telephone relays could be analysed through Boolean algebra. Shannon also showed that logical problems, such as adding two binary numbers together, could be modelled exactly using telephone relays — and, soon enough, vacuum tubes, transistors and semiconductor chips too. Having grounded modern computing in this way, Shannon created 'An Algebra for Theoretical Genetics', as his doctoral thesis was titled, in 1940.

A full-blown information theory arrived soon after. Following a year at the Institute for Advanced Study in Princeton, New Jersey, Shannon joined Bell Telephone Laboratories full-time in 1941 and worked on wartime cryptography and fire-control projects. Shrouded in secrecy, the wartime work introduced him to British computer scientist Alan Turing (they had tea together for two months but could not discuss their code work), and linked his thinking with mathematician Norbert Wiener's broadly similar theory of the role of randomness or 'entropy' in information. All communication, Shannon decided, resembled coded messages sent through a noisy channel: distortion and noise battled against redundancy and bandwidth. He defined these terms mathematically. Shannon gained international acclaim after publishing two famous technical articles in which he named the bit (short for 'binary digit', first coined in a Bell Labs memo by statistician John Tukey), and after writing a popular book with engineer

and mathematician Warren Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, 1949).

Having engagingly assembled information theory, Gleick might have examined its many ramifications in the mathematics of coding theory, data compression and error correction that underpins everything from mobile phones to DVDs. Instead, he treats information more metaphorically, covering the founding of cybernetics, the genetic code of DNA and the birth of quantum information science and the allure of quantum computing. His asides on the editing history of Wikipedia articles, although entertaining, begin to stretch the interpretive framework. A final substantive chapter surveys our predicament of information overload, the flood of the book's subtitle: too many genomes, sky surveys and climate models, let alone e-mails. He confronts social media such as Twitter, describing it as "banality shrink-wrapped, enforcing triviality by limiting all messages to 140 characters".

Gleick admirably raises the question of how information relates to meaning and semantics, which Shannon specifically ruled out of scope in his theory, yet Gleick mostly inclines towards instances of verbal and mathematical thinking. "The written word — the persistent word — was a prerequisite for conscious thought as we understand it," Gleick suggests. An anecdote from Feynman opens up this tidy world: "Thinking is nothing but talking to yourself," he once remarked. "Oh yeah?" countered a friend. "Do you know the crazy shape of the crankshaft in a car? Now tell me: how did you describe it when you were talking to yourself?" Thinking in this spatial way, Feynman set the stage for nanotechnology in his article in *Popular Science* in November 1960, titled 'There's Plenty of Room at the Bottom: How to Build an Automobile Smaller than this Dot.' Semantics and spatial thinking might be considered for a new, generative theory of information to enhance Shannon's. ■

Thomas J. Misa is director of the Charles Babbage Institute at the University of Minnesota, Minneapolis, Minnesota 55455, USA.
e-mail: tmisa@umn.edu



CHAIN (1997) BY DALTON GHETTI.
COURTESY OF THE ARTIST; PHOTO: S. HOWARD



Fermilab: Physics, the Frontier, and Megascience

Lillian Hoddeson, Adrienne W. Kolb and Catherine Westfall (Univ. Chicago Press, 2011; \$30)
For 40 years, the Fermi National Accelerator Laboratory in Illinois has stood at the frontier of high-energy physics. The book charts the rise of this institution, detailing the difficulties of balancing pioneering science with tightened budgets.



The Eerie Silence

Paul Davies (Mariner Books, 2011; \$15.95)
Astrophysicist Paul Davies describes the 50-year Search for Extra-Terrestrial Intelligence project. He proposes other approaches, from scouring Earth for microscopic aliens to seeking intelligence on planets beyond the Solar System. (See Chris McKay's review: *Nature* **464**, 34; 2010.)

PSYCHOLOGY

Holding on to happiness

Sonja Lyubomirsky welcomes a call for society to encourage people to ‘flourish’.

The premise of positive psychology — that it is as important to investigate wellness as it is to study misery — has reached the mainstream. Discussed routinely by politicians, educators and mental-health professionals, the field’s influence has grown rapidly. Martin Seligman, director of the Positive Psychology Center at the University of Pennsylvania in Philadelphia, is the scholar, educator and charismatic leader who has championed these ideas passionately for more than a decade. In *Flourish*, his most personal and boldest book so far he argues that we should set aside “happiness” as a goal, and embrace a broader measure of well-being, which he calls “flourishing”.

Seligman shares a wealth of insights and stories, mostly compelling and sometimes maddeningly digressing, which cast light on his passions and pet peeves. His wisdom and audacious opinions explain why he has attracted legions of both followers and high-profile critics, including writer and columnist Barbara Ehrenreich and *New Yorker* journalist Jane Mayer.

Two themes run through the book. The first is that the study of optimal human functioning must be grounded in rigorous science. The second is more controversial: positive-psychology researchers have a duty to make the world a better place. Seligman’s book is a paean to applied science, a blueprint for how to translate empirical evidence from the laboratory to the real world.

Seligman describes several applied initiatives that he has conceived and shepherded. In education, he has created and implemented curricula to develop character strengths (such as kindness and leadership), build grit (passion and perseverance) and enhance positive emotions (happiness and gratitude) in schoolchildren and undergraduates. For example, children at risk of depression are guided to identify their top signature strength (such as loyalty)

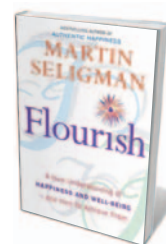
and use it in a new way at school each week. Seligman also teaches the theory and research behind positive psychology to individuals in a range of occupations — from life coaches and entrepreneurs to policy wonks and fitness instructors.

The most impressive effort Seligman discusses is the Comprehensive Soldier Fitness programme now being implemented across the US Army community. The programme, profiled in a special issue of *American Psychologist* this January, involves measuring “psychosocial fitness” and building resilience in several life domains: emotional, social, family and spiritual. For example, sergeants are trained to avoid thinking the worst when faced with adversities, and soldiers are taught to identify emotions in others. This is a rare

opportunity to change the culture of a huge institution that is not known for prioritizing emotions, to prevent suffering (including suicide and post-traumatic stress) and bolster both flourishing and effectiveness in military roles.

It is no accident that this book is titled *Flourish* yet Seligman’s preceding best-seller was called *Authentic Happiness*. He professes that he now detests the word “happiness”, for three reasons: it is overused and nearly meaningless; it is measured subjectively; and it connotes smiley-faced cheerfulness and hedonism. However, the alternative terms proffered by Seligman — flourishing, well-being, meaning, love and growth — are no more likely to elude these problems.

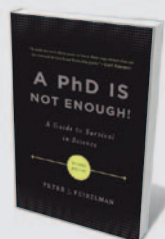
Setting out a new theory of well-being, Seligman posits that flourishing has four elements or pillars: positive emotion (happiness, satisfaction, engagement); meaning; positive relationships; and accomplishment (mastery). It is hard to argue with this intuitively appealing thesis. However, it has its weaknesses. First, Seligman’s theory confuses the elements of well-being with the contributors and consequences of well-being. For instance, people who report that they are happy are more likely than their less-satisfied peers to have meaning, good relationships and accomplishment in their lives. These factors may be sources of happiness — having a good marriage makes one more happy, for example. Or they may be outcomes — happier people are likely to forge satisfying relationships.



Flourish: A Visionary New Understanding of Happiness and Well-Being
MARTIN SELIGMAN
Free Press/Nicholas Brealey Publishing:
2011. 368 pp/408 pp.
\$26/£14.99

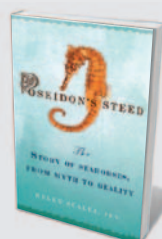


Size Matters: *Alces alces* (moose email) by Dana Harel, 2009.



A PhD is Not Enough! A Guide to Survival in Science
Peter J. Feibelman (Basic Books, 2011; \$14.95)

Climbing the scientific career ladder is difficult, and the first steps from doctoral student to postdoc are the most precarious. Drawing on his experience as a physicist in academic and government labs, in his new edition Peter Feibelman offers career guidance to those entering the research job market.



Poseidon's Steed: The Story of Seahorses, from Myth to Reality

Helen Scales (Gotham Books, 2010; \$15)
The weird world of the seahorse is explored by marine biologist Helen Scales. She describes its peculiar biology and the threats it faces, and reveals its importance to humans, from its role in Chinese medicine to ancient seahorse myths.

COURTESY OF THE ARTIST/FREY NORRIS GALLERY, SAN FRANCISCO

Second, although the four pillars are correlated, they do not necessarily amount to a single structure: they need not occur together and may originate and evolve differently over time. For example, a public servant who is passionate about his work may experience a great deal of positive emotion and meaning over the course of his career, but his relationships may suffer if he overworks. A selfless Mother-Theresa type may have meaning, accomplishment and fulfilling relationships, but experience little joy. These examples point to a third concern: it is not clear whether Seligman's conceptualization of well-being is shared among cultures.

Third, there is no empirical evidence that constructs such as meaning or love can be measured more objectively than happiness. If happiness is "all in one's head", as Seligman asserts, then so are some of the four pillars. Terms such as flourishing and well-being are useful shorthand, but calling the four pillars a theory is premature.

Seligman's ideas have a great deal of merit, but it is too soon to dispense with happiness. Research reveals that happy people are not self-centred, gratification-seeking hedonists whose lives are lacking in meaning or fulfilment. On the contrary, hundreds of studies have shown that happiness relates to outcomes such as creativity, productivity, effective coping, satisfying marriages, close friendships, higher earnings, longevity and strong immune systems.

Seligman's galvanizing goal for positive psychology is for 51% of the world's population to be flourishing by the year 2051. Unlike many authors, he offers detailed and tested solutions as well as compelling arguments for how societies can aim to raise the amount of positive emotion, meaning, good relationships and accomplishment in their citizens. Even if his four pillars don't quite make a theory, everyone stands to benefit from his initiatives. If they are happy, flourishing or enjoying well-being, people won't care about the labels that researchers attach to those good feelings. ■

Sonja Lyubomirsky is professor of psychology at the University of California, Riverside, California 92521, USA, and author of *The How of Happiness*. e-mail: sonja.lyubomirsky@ucr.edu



Size Matters: Second Growth Forest (2008) by Eamon MacMahon.

EARTH SYSTEMS

The biosphere rebooted

Michael J. Benton finds hope for the future in a study of humanity's cooperation with the environment.

Many recent books about the fate of life on Earth muse on fragility, tipping points and crises. But some writers see a more hopeful future for the planet. Without ignoring the monumental threats posed by humans, interdisciplinary studies may be offering reasons to be cheerful about the resilience of life in the face of change, and our chances of surviving this and the next century. Australian palaeontologist Tim Flannery's *Here on Earth* follows in this optimistic vein.

By tracing the great shifts in Earth's geochemical and biological systems through time, he argues that life generates ever-more-sophisticated responses to varying planetary conditions. In particular, he notes, "from the most intense competition for survival, cooperation has emerged". Such natural transformations hold lessons



Here on Earth: A New Beginning

TIM FLANNERY
Allen Lane/Atlantic
Monthly Press: 2011.
336 pp/288 pp.
£14.99/\$25

for future challenges. He develops his theme through parallel accounts of the history of Earth and of life, harnessing an impressive mix of research in geology, chemistry, biology, palaeoanthropology and sociology.

Flannery moves deftly through some difficult science. Early in the book, he espouses British environmentalist and chemist James Lovelock's Gaia hypothesis that life stabilizes the planet and makes it habitable. He explains how chemical cycling during

COURTESY OF THE ARTIST



About a Mountain

John D'Agata (W. W. Norton, 2011; \$14.95)
Writer John D'Agata investigates the US government's plan to store nuclear waste beneath Yucca Mountain in Nevada. He documents the history of the project, its supporters and detractors, and muses on atomic-bomb tests and Las Vegas's diminishing water supply and high suicide rate.



Nature's Palette: The Science of Plant Color

David Lee (Univ. Chicago Press, 2010; \$22.50)
The science of plant colour is explored by botanist David Lee, from the decorative use of plant dyes to the chemistry of plant leaf colour. "A compelling case that botany is full of intellectual challenges, many shamefully neglected," wrote Philip Ball in his review (*Nature* **449**, 982; 2007).

Second, although the four pillars are correlated, they do not necessarily amount to a single structure: they need not occur together and may originate and evolve differently over time. For example, a public servant who is passionate about his work may experience a great deal of positive emotion and meaning over the course of his career, but his relationships may suffer if he overworks. A selfless Mother-Theresa type may have meaning, accomplishment and fulfilling relationships, but experience little joy. These examples point to a third concern: it is not clear whether Seligman's conceptualization of well-being is shared among cultures.

Third, there is no empirical evidence that constructs such as meaning or love can be measured more objectively than happiness. If happiness is "all in one's head", as Seligman asserts, then so are some of the four pillars. Terms such as flourishing and well-being are useful shorthand, but calling the four pillars a theory is premature.

Seligman's ideas have a great deal of merit, but it is too soon to dispense with happiness. Research reveals that happy people are not self-centred, gratification-seeking hedonists whose lives are lacking in meaning or fulfilment. On the contrary, hundreds of studies have shown that happiness relates to outcomes such as creativity, productivity, effective coping, satisfying marriages, close friendships, higher earnings, longevity and strong immune systems.

Seligman's galvanizing goal for positive psychology is for 51% of the world's population to be flourishing by the year 2051. Unlike many authors, he offers detailed and tested solutions as well as compelling arguments for how societies can aim to raise the amount of positive emotion, meaning, good relationships and accomplishment in their citizens. Even if his four pillars don't quite make a theory, everyone stands to benefit from his initiatives. If they are happy, flourishing or enjoying well-being, people won't care about the labels that researchers attach to those good feelings. ■

Sonja Lyubomirsky is professor of psychology at the University of California, Riverside, California 92521, USA, and author of *The How of Happiness*. e-mail: sonja.lyubomirsky@ucr.edu



Size Matters: Second Growth Forest (2008) by Eamon MacMahon.

EARTH SYSTEMS

The biosphere rebooted

Michael J. Benton finds hope for the future in a study of humanity's cooperation with the environment.

Many recent books about the fate of life on Earth muse on fragility, tipping points and crises. But some writers see a more hopeful future for the planet. Without ignoring the monumental threats posed by humans, interdisciplinary studies may be offering reasons to be cheerful about the resilience of life in the face of change, and our chances of surviving this and the next century. Australian palaeontologist Tim Flannery's *Here on Earth* follows in this optimistic vein.

By tracing the great shifts in Earth's geochemical and biological systems through time, he argues that life generates ever-more-sophisticated responses to varying planetary conditions. In particular, he notes, "from the most intense competition for survival, cooperation has emerged". Such natural transformations hold lessons



Here on Earth: A New Beginning

TIM FLANNERY
Allen Lane/Atlantic
Monthly Press: 2011.
336 pp/288 pp.
£14.99/\$25

for future challenges. He develops his theme through parallel accounts of the history of Earth and of life, harnessing an impressive mix of research in geology, chemistry, biology, palaeoanthropology and sociology.

Flannery moves deftly through some difficult science. Early in the book, he espouses British environmentalist and chemist James Lovelock's Gaia hypothesis that life stabilizes the planet and makes it habitable. He explains how chemical cycling during

COURTESY OF THE ARTIST



About a Mountain

John D'Agata (W. W. Norton, 2011; \$14.95)
Writer John D'Agata investigates the US government's plan to store nuclear waste beneath Yucca Mountain in Nevada. He documents the history of the project, its supporters and detractors, and muses on atomic-bomb tests and Las Vegas's diminishing water supply and high suicide rate.



Nature's Palette: The Science of Plant Color

David Lee (Univ. Chicago Press, 2010; \$22.50)
The science of plant colour is explored by botanist David Lee, from the decorative use of plant dyes to the chemistry of plant leaf colour. "A compelling case that botany is full of intellectual challenges, many shamefully neglected," wrote Philip Ball in his review (*Nature* **449**, 982; 2007).

the Precambrian era — the first 4 billion years of Earth's existence, until 542 million years ago — led animals, including humans, to develop the ability to absorb and store poisonous elements such as mercury, cadmium and lead. He explores how the early evolution of life built the atmosphere; and how continental drift and mid-ocean hydrothermal vents known as black smokers maintain the salinity of the sea for marine life.

Flannery then switches to human evolution and migration through Australia, Asia, Europe and the Americas, focusing on human ancestors' impacts on the land. He shows, for example, how the slaughter of mammoths in the Siberian tundra effectively destroyed the productivity of this terrain. Tundra plants must be eaten for the carbon they contain to be recycled, otherwise they simply freeze and the nutrients are locked in. Mammoths were the greatest eaters of this modest plant cover, bulldozing the snow aside with their baroque tusks and redepositing the digested remnants as copious urine and droppings, which fertilized the land. With the demise of the mammoths, the tundra's productivity also declined.

Yet human behaviour in prehistory, and in non-industrialized societies today, was not always environmentally destructive. Flannery relates how Australian Aboriginal people learned that nutrients were recycled when vegetation was burnt in small patches — in contrast to the vast interior deserts created by mechanized agriculture on the continent today. Through taboos over eating certain rare species, indigenous New Guineans effectively preserved local biodiversity.

Central to Flannery's optimism is cooperation, including that between humans and the environment. He looks to insect colonies — sometimes termed 'superorganisms' because they can act as one unit — in which individual members follow pheromone messages to

fulfil tasks that meet group objectives. Comparing insect communities to human societies, Flannery shows how cooperation has increased the lifespan and benefits of modern humans compared with our ancestors, who — although they were able to tackle almost any task — had short and painful lives.

Flannery acknowledges our persistent efforts to destroy Earth and ourselves:

THE SLAUGHTER OF MAMMOTHS IN THE SIBERIAN TUNDRA EFFECTIVELY DESTROYED THE PRODUCTIVITY OF THIS TERRAIN.



Size Matters: *William River* (2008) by Eamon MacMahon.

nuclear proliferation, agricultural spraying with toxic weed killers and insecticides, persistent organic pollutants and industrial effluent of metals and carbon dioxide. The narrative in each case of big business, disease and death, research and eventual regulation has been told many times, but rarely as thoroughly and dispassionately. The biggest threat of all, Flannery contends, is overpopulation.

Yet he agrees with United Nations estimates that humans will self-regulate at about

9 billion individuals by 2050, thanks to rising standards of living and decreasing family size. Improving economies will also strengthen people's reasons to invest in their future. As individuals and corporations stop "discounting the future" by taking a reckless view of their own and their community's survival, they will adopt more sustainable lifestyles, in which conspicuous consumption is mocked

rather than admired. In support, Flannery notes the rise in the number of democratic countries from 40 to 123 in the past 50 years, the beginning of international negotiations about sustainability and the rise of the Internet and mobile phones, which make secrecy hard

to maintain.

Despite the tendency for people to ignore the inevitable, and to become immune to doom-laden prophecies, Flannery believes that humanity will act before New York, Shanghai and London sink beneath the waves. A combination of effective recycling of carbon back into the soil, rewilding of vast areas and natural stabilization of human populations could present a long-term model for survival. It is a clear and rational proposal. However, many futurists would deny his optimism: reversing levels of current carbon usage, for example, would require an unimaginable change in cooperative behaviour worldwide.

Although some might quibble about his reliance on speculative concepts such as Gaia, Flannery's command of evolution, environmental chemistry, civilization and human motivations strengthens his case. His buoyant futurology is a hopeful counterpoint to the short-term denial and inertia of so many current decision-makers. ■

Michael J. Benton is professor of vertebrate palaeontology at the University of Bristol, Bristol BS8 1RJ, UK.
e-mail: mike.benton@bristol.ac.uk



Living at Micro Scale: The Unexpected Physics of Being Small

David B. Dusenbery (Harvard Univ. Press, 2011; \$22.95)
The size, shape and behaviour of tiny organisms are challenged and constrained by physics. Biologist David Dusenbery describes how factors that larger organisms can ignore — such as the viscosity of water or air — affect microorganisms.



Diversity and Complexity

Scott E. Page (Princeton Univ. Press, 2010; \$19.95)
Complex systems respond to diversity in sophisticated ways — some of which enhance system performance. Theorist Scott Page explains how diversity affects biological, ecological and social systems from tropical environments to the economy.

CORRESPONDENCE

Intolerance: retain healthy scepticism

The UK government's chief scientist, John Beddington, has done much to promote the potential of science to address pressing global imperatives. But in a straight-talking speech last month, he urged his audience of 300 government scientists to be "grossly intolerant" of "pernicious" and "fatuous" "pseudoscience". In this he included: scepticism of genetic modification technology; "illegitimate" advocacy of environmental precaution in response to unknowns; and suggestions that science is subject to morality. This approach is a rejection not just of irrational denial but of entirely reasonable social scepticism concerning science itself.

It is ironic that defence of scientific rationality is often so emotive. Science does not monopolize social rationality. Although imperfectly realized, the principal distinctions between science and, say religion, politics, commerce or the media are arguably social practices of organized scepticism. Open publication, peer review, experimentation and critical respect for evidence help promote reasoned argument. But rational scepticism is as important outside as inside the social practices of science. Hence the motto of Britain's Royal Society, '*Nullius in verba*': take nothing solely on authority — even from scientists.

Suppression of rational scepticism of science is also potentially undemocratic. Allocation of research priorities depends on contending interests and values. Technologies are partly shaped by social and political factors, not simply determined by unfolding knowledge. Economic, institutional, military and political pressures help to drive

world science in particular directions. The internal dynamics of disciplines and universities are not immune to personality, privilege, prejudice or power. Intrinsic scientific uncertainties underscore the role of democratic debate.

Inhibiting reasonable social questioning of science can foster disingenuity, polarization and untrustworthiness. The progressive social potential of science is therefore supported, not hindered, by greater political tolerance for scepticism.

Andy Stirling *University of Sussex, UK.*
a.c.stirling@sussex.ac.uk

Intolerance: science informs, not defines

The UK government's chief scientist, John Beddington, last month demanded intolerance towards the "pseudoscience" used to challenge government science on certain policy issues. However, none of the growing range of public issues involving important scientific questions can be reduced, as Beddington did, simply to "science" or "pseudoscience".

The UK government's scientific advisory apparatus routinely imposes this 'scientism' — to science's own public detriment.

Public disagreement with policy commitments such as genetically modified crops or vaccination of children against measles, mumps and rubella (MMR) is rarely based on opposition to the science itself, even when science is used exclusively to justify policy. People are more likely to object to the brushing aside of their reasonable questions by scientists and policy-makers, and to the exclusion of important factors such as the pervasive exaggeration of scientific understanding and predictive control.

Take the now-discredited claim that the MMR vaccine

causes autism: for dissenting parents, the question was whether their child might be at greater individual risk than the average-population risk figures used by the government. This separate question was apparently discounted by government and supporting scientists. Yet which question was more salient was a matter of legitimate difference, not one-dimensional science versus pseudoscience. Both required recognition.

What policy advisers anoint as 'science' for intended public authority always embodies unstated policy-related commitments, including presumptions over the defining questions. Such social questions in public science should be recognized and debated openly. Scientific knowledge should inform public issues, not define them.

Brian Wynne *ESRC Centre for Economic and Social Aspects of Genomics (Cesagen), Lancaster University, UK.*
b.wynne@lancaster.ac.uk

More protection for China's wetlands

We estimate that China's natural wetlands are disappearing even faster than feared (*Nature* 471, 19; 2011). Stricter measures are needed to protect what is left of this valuable ecosystem against the increasing demand for land and development.

Using Landsat Thematic Mapper imagery to map China's wetlands (see, for example, *Nature* 458, 134; 2009), we found that 33% were lost between 1978 and 2008. Some 55% of these were natural inland marshes, many of which are biodiversity hotspots. Land reclamation accounted for more than 70% of the total loss.

The Tibetan plateau generated about 6,000 square kilometres of new wetlands between 1990 and 2008 through deglaciation

and thawing of permafrost, reducing wetland losses from 66% between 1978 and 1990 to just 6% in 2000–08. Matters are also improving as a result of the Chinese government's substantial efforts in creating new protection areas and initiating wetland restoration projects.

However, more effort is needed to arrest this loss altogether. This should be directed at providing operational support for continuous monitoring of critical wetlands, integrating wetland protection into watershed management plans and specific legislation for wetland protection. **Zhenguo Niu, Haiying Zhang** *State Key Lab of Remote Sensing Science, Institute of Remote Sensing Applications, Beijing, China.* **Peng Gong** *University of California, Berkeley, USA.*
penggong@berkeley.edu

The social impact of innovation

David Edgerton wants scientists to be more vocal in opposing hare-brained innovations that stand little or no chance of success (*Nature* 471, 27–29; 2011).

I agree that we place too high a value on innovation for its own sake, but fail to see what this has to do with Luddism. The objection of the Luddites was not to money being wasted on machines that would not work. On the contrary, the machines that the Luddites tried to destroy worked only too well.

The lesson to be learned from Luddism is both more important and more sophisticated: when evaluating the worth of an innovation, we should be concerned with its wider social implications, as well as its narrowly economic ones. In other words, scientists should have consciences.

Timothy Roper *University of Sussex, UK.*
t.j.roper@sussex.ac.uk

Marshall Stoneham

(1940–2011)

Theoretician who contributed widely to condensed-matter physics.

Marshall Stoneham was a theoretical physicist, most noted for his work on defects in solids, who made wide-ranging contributions to condensed-matter science. He was a leading figure in the Atomic Energy Research Establishment (AERE) near Harwell, UK, when the key challenges to the nuclear industry became materials and waste disposal as much as nuclear physics. He died on 18 February, aged 70.

Born in 1940 in Barrow-in-Furness, in northern England, Stoneham was educated at Barrow Grammar School for Boys, which produced three future Fellows of the Royal Society (of whom he was one), all in engineering and physical sciences, in the space of 15 years. He studied physics at the University of Bristol and remained there to do a PhD, under the supervision of Maurice Pryce, in the theory of spin-lattice relaxation — how the spins of ions in a solid interact with the host crystal lattice — a topic of current importance in quantum information.

In 1964 he joined the Theoretical Physics Division at AERE Harwell, then one of the world's foremost physical-sciences laboratories. The division, built up by Walter Marshall and Alan Lidiard, included leading figures such as Ron Bullough (in the theory of elasticity in metals), Tony Lane (in nuclear theory) and John Hubbard (in strongly correlated electron physics). Its work was stimulated by the needs of the nuclear industry, and Stoneham flourished in this challenging environment.

He worked mainly on defects in solids, because these determined many properties of materials important in nuclear applications and in the emerging semiconductor industry. He was particularly known for explaining, with Pete Flynn at the University of Illinois at Urbana-Champaign, how hydrogen atoms move rapidly through metals and hence contribute to brittleness. This was typical of his style, combining a sharp fundamental insight (how the diffusing quantum atom interacts with the thermal environment) with the solution to an important applied problem.

In 1975, he published the influential monograph *Theory of Defects in Solids*, which rationalized the vast phenomenology of the subject and explained it in terms of the underlying quantum mechanics, giving the field an interpretative tool that is still valuable today. Collaborators who came to know him in subsequent years were struck by his astonishing memory — he often answered questions in

the form of a page reference to the book!

During the 1970s, the Harwell theory team had a leading role in the development of condensed-matter computational physics, enabled by the availability of mainframe computers. Although Stoneham was at heart an analytical theoretician, he appreciated this new capability. His solid-state and quantum-physics group exploited Harwell's HADES code to model defects in solids, including their formation and migration energies and key structural properties. The influence of this code, and Stoneham's group, extended beyond defect physics into



solid-state chemistry, especially the chemistry of non-stoichiometric compounds — the component elements of which are not present in simple proportions, usually indicating the presence of defects. Indeed, the group became a magnet in the 1970s and 1980s for those working on defects in materials, and the work extended to surfaces and interfaces, and made key predictions of the structures and stabilities of oxide surfaces.

Stoneham's group made advances in the physics and chemistry of nuclear fuels and, as the remit of Harwell broadened, in a wider range of problems. One such was Tony Harker's work on non-destructive inspection of gas pipelines using ultrasound. Stoneham embraced the increasing commercial emphasis of what was by then AEA Industrial Technology, and in the early 1990s became AEA's chief scientist. Throughout his career he was a leading expert on, and exponent of, fission nuclear power. Just last year

he wrote a masterly article (A. M. Stoneham *Phil. Trans. R. Soc. A* **368**, 3295–3313; 2010) on the history of the UK nuclear-energy programme and its prospects and challenges.

Nevertheless, with the environment at Harwell becoming less conducive to fundamental research, in 1995 Stoneham became the first Massey professor of physics at University College London and director of its Centre for Materials Research. He loved the wide range of materials-related work there and developed projects in areas as diverse as minimally invasive dentistry, odour recognition, diamond film growth and quantum information science, where his ideas on optically controlled gates led to a substantial and ongoing research programme.

Stoneham was a major figure in the UK and international physics communities, with awards to match. He was for several years editor of *Journal of Physics: Condensed Matter* and last year became president of the Institute of Physics. He also inspired the many scientists who worked with him and guided them at crucial stages in their careers. He was exceptionally supportive of his junior collaborators, and enjoyed challenging them to solve problems, using modern computational methods, more quickly and accurately than he could with his ingenuity and his slide rule.

Marshall's wife Doreen and two daughters are all physical scientists — a fact in which he took considerable pride, and with Doreen he founded Oxford Authentication, a company that uses thermoluminescence to date fine-art ceramics.

His scientific range was broad, but his interests extended even further. He was an accomplished horn player and the author of a book on music for wind ensembles. He was always stimulating company and he seemed to know something interesting about — and held a strong opinion on — almost any subject. He gained and retained the respect and affection of all who worked with him and he leaves behind an enduring contribution to physics and to physicists. ■

Richard Catlow is dean of mathematical and physical sciences at University College London, London WC1H 0AJ, UK. He knew Stoneham since doing a PhD at Harwell in the 1970s. **Andrew Fisher** is professor of physics at University College London, London WC1E 6BT, UK. He became Stoneham's PhD student in 1986. e-mail: c.r.a.catlow@ucl.ac.uk

VISION

Dicer leaps into view

An enzyme called Dicer is known for its central role in RNA-controlled gene silencing. Mammalian Dicer1, however, also seems to have another function: it maintains visual health by degrading toxic RNA molecules. [SEE LETTER P.325](#)

GUNTER MEISTER

As their name suggests, Dicer enzymes cut long double-stranded RNA molecules into shorter pieces^{1,2}. They are therefore crucial for gene-silencing pathways that involve small RNAs such as short interfering RNAs or microRNAs — the most abundant classes of small RNAs in mammals^{3,4}. Without functional Dicer, most of these RNAs cannot be generated⁵. On page 325 of this issue, Kaneko *et al.*⁶ uncover a function of mammalian Dicer that is independent of small-RNA generation but apparently equally essential.

The authors were interested in the molecular basis of geographic atrophy — an advanced form of an eye disease called age-related macular degeneration, which is a common cause of blindness in industrialized countries⁷. They found that patients with this condition have reduced levels of Dicer in their retinal pigment epithelium (RPE), an eye-specific tissue that is affected in geographic atrophy. Kaneko and colleagues also noted that in mice lacking Dicer selectively in the RPE, this tissue is severely degenerated, similar to the situation in geographic atrophy, and therefore further implicating reduced Dicer levels in the pathogenesis of the disorder (Fig. 1).

Loss of Dicer intuitively points to a role for microRNAs (miRNAs) in geographic atrophy, especially given that these short regulatory RNA sequences are highly abundant in mammalian cells. Surprisingly, however, Kaneko *et al.*⁶ could find no evidence to support this. For example, mice lacking two other enzymes required for miRNA processing (Drosha and DGCR8) did not show degeneration of the RPE, even though these animals showed severely impaired miRNA function. So how does Dicer deficiency contribute to geographic atrophy? Kaneko and colleagues describe an as-yet-unrecognized function of the enzyme.

Dicer processes double-stranded RNA, and the authors speculated that a class of double-stranded RNA other than miRNAs might be involved in the pathology of geographic

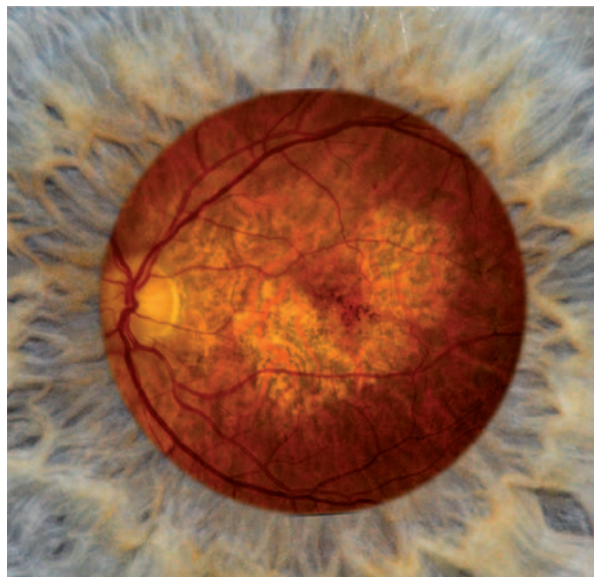


Figure 1 | Geographic atrophy. The retina of a person with geographic atrophy — an advanced age-related macular degeneration — projected through the pupil and surrounded by the iris. The retina shows a horseshoe-shaped area where the retinal pigment epithelium cells are atrophied. This creates what is known as a window defect, making visible the choroidal vessels, which the healthy retinal pigment epithelium would otherwise mask. Kaneko *et al.*⁶ implicate Dicer in the pathogenesis of geographic atrophy.

atrophy. Indeed, they detected a marked increase in the levels of *Alu* RNA sequences in the eyes — and specifically in the RPE — of patients with this condition, but not in those of healthy individuals. And they noted that mice with Dicer deficiency in the RPE express increased levels of *Alu* RNAs.

Alu RNAs are transcripts of *Alu* elements — the most common non-coding, repetitive DNA sequences in the human genome. But how do they cause geographic atrophy? Kaneko *et al.* tackled this question from several angles.

Both enhancing the expression of *Alu* RNAs and their direct injection into the eyes of normal mice reduced cell viability in the RPE and led to degeneration of this tissue. But when the authors increased the expression of Dicer along with that of *Alu* RNAs, the toxic effects of *Alu* RNAs were no longer seen. This suggests that Dicer processes longer *Alu* RNAs into shorter, non-toxic sequences. To test this possibility, the authors injected long double-stranded *Alu*

RNA, before and after treatment with Dicer, into the eyes of normal mice. Only the long unprocessed *Alu* RNAs caused RPE degeneration.

Kaneko *et al.*⁶ propose the following model for the pathogenesis of geographic atrophy. In the RPE, long double-stranded *Alu* RNA is constantly produced. In healthy individuals, Dicer processes these toxic RNAs into shorter, non-toxic versions. In patients with geographic atrophy, however, Dicer levels are greatly reduced, allowing the long *Alu* RNAs to cause cell death and RPE degeneration.

As Dicer normally produces functional classes of small RNAs from double-stranded RNA precursors, an intriguing question is whether the products of *Alu* RNA processing by Dicer have specific cellular functions. In support of this possibility, Dicer generates endogenous short interfering RNAs from repetitive elements in the mouse germ line⁸. Therefore, it could be that, in addition to the toxic long *Alu* RNAs, the absence of Dicer-generated small *Alu* RNAs contributes to features of geographic atrophy.

Previous work⁹ has shown that reduced Dicer levels can occur in many tissues and are associated with various diseases. But these studies did not analyse *Alu* RNAs. On the basis of Kaneko and co-workers' data, it is conceivable that the effects of toxic *Alu* RNAs are also widespread, as has always been expected. It remains unclear, however, why Dicer expression is reduced in the RPE, and not in other tissues. Are tissue-specific transcription factors also missing in geographic atrophy, or is Dicer degraded in the RPE after synthesis?

Another pertinent question is whether these results⁶ could further the search for new strategies for treating geographic atrophy. Kaneko *et al.* report that injecting antisense oligonucleotides into mouse eyes, which reduces *Alu* RNA levels, also decreases RPE degeneration. These are heartening findings. But before any therapy development can be contemplated, the molecular basis of *Alu* RNA detoxification by antisense oligonucleotide must be

MICHAEL HANSON & BRADLEY GELFAND

elucidated. Strategies to increase Dicer expression might also prove to be useful therapeutic approaches. ■

Gunter Meister is in the Department of Biochemistry I, Faculty of Biology and Preclinical Medicine, University of Regensburg, 93053 Regensburg, Germany. e-mail: gunter.meister@vkl.uni-regensburg.de

1. Jinek, M. & Doudna, J. A. *Nature* **457**, 405–412 (2009).

2. Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. *Cell* **118**, 57–68 (2004).
3. Krol, J., Loedige, I. & Filipowicz, W. *Nature Rev. Genet.* **11**, 597–610 (2010).
4. Carthew, R. W. & Sontheimer, E. J. *Cell* **136**, 642–655 (2009).
5. Bartel, D. P. *Cell* **136**, 215–233 (2009).
6. Kaneko, H. *et al.* *Nature* **471**, 325–330 (2011).
7. Ambati, J., Ambati, B. K., Yoo, S. H., Ianchulev, S. & Adamis, A. P. *Surv. Ophthalmol.* **48**, 257–293 (2003).
8. Kim, V. N., Han, J. & Siomi, M. C. *Nature Rev. Mol. Cell Biol.* **10**, 126–139 (2009).
9. Lambert, I. *et al.* *Cell Death Differ.* **17**, 633–641 (2010).

MATERIALS SCIENCE

Complex order in soft matter

Spherical micelles can aggregate into highly organized structures. New micelle arrangements mimic known atomic crystals, both periodic and aperiodic, and provide evidence for a material with 18-fold rotational symmetry.

SHARON C. GLOTZER & MICHAEL ENGEL

Crystallization as an ordering phenomenon is not restricted to atoms. Molecules, polymers, nanometre-sized particles and colloids also form crystals, following similar thermodynamic principles. This realization has led to considerable efforts to mimic atomic crystals on nanometre and larger scales to obtain novel materials that can, for example, trap and bend light in unusual ways. One class of structure that has captured the attention of soft-matter scientists is aperiodic crystals known as quasicrystals, whose rotational symmetries were once thought to be forbidden in solids. In work published in the *Proceedings of the National Academy of Sciences*, Fischer *et al.*¹ report soft-matter quasicrystals with 12-fold and, for the first time, 18-fold rotational symmetry. The structures assemble from spherical micelles that self-organize from commonplace block copolymers in water.

A micelle can be roughly envisaged as a stiff core surrounded by a brush-like corona of flexible tethers. Micelles form readily in solutions of amphiphiles (molecules with solvent-loving and solvent-hating parts, which include surfactants, block copolymers and branched polymers known as dendrimers) and are easily deformable under thermal motion. Thermodynamic-equilibrium phases are characterized by a competition between elastic forces favouring spherical shapes and forces that tend to minimize the contact area between neighbouring micelles.

Although a complete classification of spherical micelle phases is an open problem, their

phase behaviour is well understood in two limiting cases. If the micelles are very soft, elastic forces can be neglected and the micelles behave like soap froths². In this case, the stable crystal is the solution of a classical problem in mathematics, attributed to Kelvin: what space-filling arrangement of identical cells of equal volume has minimal surface area? The optimal solution is the Weaire–Phelan foam, the geometric dual of the A15 crystal structure,

which consists of micelles located at the corners and in the centre of a cubic unit cell, and arranged in pairs on the cube faces. A close, but inferior competitor is the simple body-centred cubic (b.c.c.) lattice, consisting of cubic cells with micelles at each vertex and one in the middle of the cell. If, on the other hand, the micelles are very stiff, they pack densely like hard spheres and arrange into the face-centred cubic lattice (f.c.c.; Fig. 1a), just like apples and oranges stacked on store shelves. Between these two limits, surprisingly complex crystals and quasicrystals are observed.

Quasicrystals were originally found only in metallic alloys. But the past decade has witnessed the discovery of quasicrystals and structurally related crystals in several soft-matter systems. The first of these was discovered³ in 2004 in a solution of wedge-shaped macromolecules (dendrons) that form micelles with radii of about 10 nanometres. A structurally closely related periodic crystal called the σ -phase was observed under similar conditions. Both phases were also observed in mixtures of iron oxide (Fe₂O₃) and gold (Au) nanoparticles⁴, and in computer simulations of hard tetrahedra⁵. Star-shaped polymers were also shown to arrange into columnar quasicrystals⁶. Block copolymers form micelles larger than dendrimer micelles and crystallize into the σ -phase⁷, and now, as Fischer *et al.*¹ demonstrate in an aqueous solution of poly(isoprene-*b*-ethylene oxide), they self-assemble into quasicrystals.

A material can obey symmetries that preserve the original structure upon rotation. Periodic order in two and three dimensions can have only 2-, 3-, 4- and 6-fold rotational

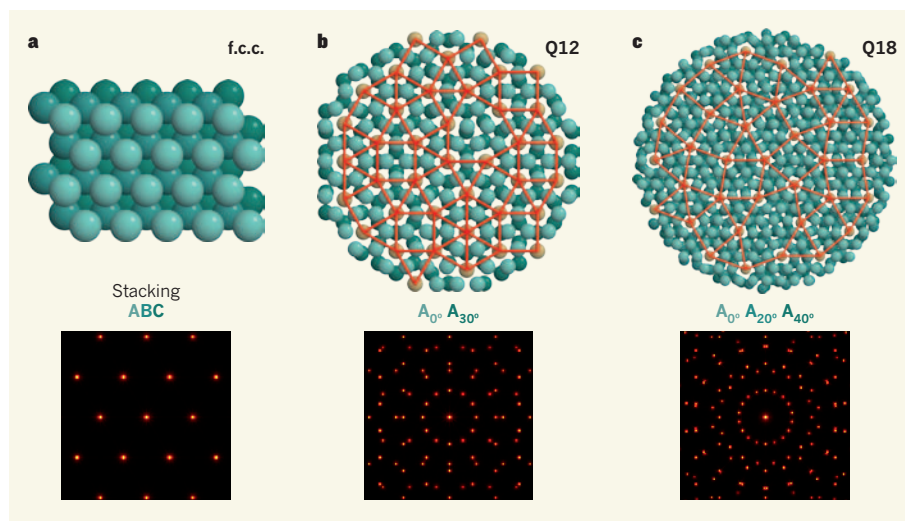


Figure 1 | Periodic and aperiodic micellar phases. **a**, Periodic face-centred cubic crystal phase (f.c.c.). **b, c**, Aperiodic quasicrystalline phases with 12-fold and 18-fold rotational symmetries (Q12, Q18). Each of these three phases can be described as a stacking of hexagonal layers of micelles (turquoise in **a**, and turquoise and yellow in **b** and **c**). In f.c.c., three flat layers A, B, C are shifted. Fischer *et al.*¹ model Q12 as two, on average hexagonal, layers that are rotated by 30° relative to each other (A₀, A₃₀), and Q18 as three layers that are rotated by 0°, 20° and 40° (A₀, A₂₀, A₄₀). The layers are not flat, and instead, the yellow micelles are positioned in between layer pairs. Red bonds indicate an exemplary tiling, and simulated diffraction patterns (bottom) indicate the phases' rotational symmetry.

M. ENGEL

elucidated. Strategies to increase Dicer expression might also prove to be useful therapeutic approaches. ■

Gunter Meister is in the Department of Biochemistry I, Faculty of Biology and Preclinical Medicine, University of Regensburg, 93053 Regensburg, Germany. e-mail: gunter.meister@vkl.uni-regensburg.de

1. Jinek, M. & Doudna, J. A. *Nature* **457**, 405–412 (2009).

2. Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. *Cell* **118**, 57–68 (2004).
3. Krol, J., Loedige, I. & Filipowicz, W. *Nature Rev. Genet.* **11**, 597–610 (2010).
4. Carthew, R. W. & Sontheimer, E. J. *Cell* **136**, 642–655 (2009).
5. Bartel, D. P. *Cell* **136**, 215–233 (2009).
6. Kaneko, H. *et al.* *Nature* **471**, 325–330 (2011).
7. Ambati, J., Ambati, B. K., Yoo, S. H., Ianchulev, S. & Adamis, A. P. *Surv. Ophthalmol.* **48**, 257–293 (2003).
8. Kim, V. N., Han, J. & Siomi, M. C. *Nature Rev. Mol. Cell Biol.* **10**, 126–139 (2009).
9. Lambert, I. *et al.* *Cell Death Differ.* **17**, 633–641 (2010).

MATERIALS SCIENCE

Complex order in soft matter

Spherical micelles can aggregate into highly organized structures. New micelle arrangements mimic known atomic crystals, both periodic and aperiodic, and provide evidence for a material with 18-fold rotational symmetry.

SHARON C. GLOTZER & MICHAEL ENGEL

Crystallization as an ordering phenomenon is not restricted to atoms. Molecules, polymers, nanometre-sized particles and colloids also form crystals, following similar thermodynamic principles. This realization has led to considerable efforts to mimic atomic crystals on nanometre and larger scales to obtain novel materials that can, for example, trap and bend light in unusual ways. One class of structure that has captured the attention of soft-matter scientists is aperiodic crystals known as quasicrystals, whose rotational symmetries were once thought to be forbidden in solids. In work published in the *Proceedings of the National Academy of Sciences*, Fischer *et al.*¹ report soft-matter quasicrystals with 12-fold and, for the first time, 18-fold rotational symmetry. The structures assemble from spherical micelles that self-organize from commonplace block copolymers in water.

A micelle can be roughly envisaged as a stiff core surrounded by a brush-like corona of flexible tethers. Micelles form readily in solutions of amphiphiles (molecules with solvent-loving and solvent-hating parts, which include surfactants, block copolymers and branched polymers known as dendrimers) and are easily deformable under thermal motion. Thermodynamic-equilibrium phases are characterized by a competition between elastic forces favouring spherical shapes and forces that tend to minimize the contact area between neighbouring micelles.

Although a complete classification of spherical micelle phases is an open problem, their

phase behaviour is well understood in two limiting cases. If the micelles are very soft, elastic forces can be neglected and the micelles behave like soap froths². In this case, the stable crystal is the solution of a classical problem in mathematics, attributed to Kelvin: what space-filling arrangement of identical cells of equal volume has minimal surface area? The optimal solution is the Weaire–Phelan foam, the geometric dual of the A15 crystal structure,

which consists of micelles located at the corners and in the centre of a cubic unit cell, and arranged in pairs on the cube faces. A close, but inferior competitor is the simple body-centred cubic (b.c.c.) lattice, consisting of cubic cells with micelles at each vertex and one in the middle of the cell. If, on the other hand, the micelles are very stiff, they pack densely like hard spheres and arrange into the face-centred cubic lattice (f.c.c.; Fig. 1a), just like apples and oranges stacked on store shelves. Between these two limits, surprisingly complex crystals and quasicrystals are observed.

Quasicrystals were originally found only in metallic alloys. But the past decade has witnessed the discovery of quasicrystals and structurally related crystals in several soft-matter systems. The first of these was discovered³ in 2004 in a solution of wedge-shaped macromolecules (dendrons) that form micelles with radii of about 10 nanometres. A structurally closely related periodic crystal called the σ -phase was observed under similar conditions. Both phases were also observed in mixtures of iron oxide (Fe_2O_3) and gold (Au) nanoparticles⁴, and in computer simulations of hard tetrahedra⁵. Star-shaped polymers were also shown to arrange into columnar quasicrystals⁶. Block copolymers form micelles larger than dendrimer micelles and crystallize into the σ -phase⁷, and now, as Fischer *et al.*¹ demonstrate in an aqueous solution of poly(isoprene-*b*-ethylene oxide), they self-assemble into quasicrystals.

A material can obey symmetries that preserve the original structure upon rotation. Periodic order in two and three dimensions can have only 2-, 3-, 4- and 6-fold rotational

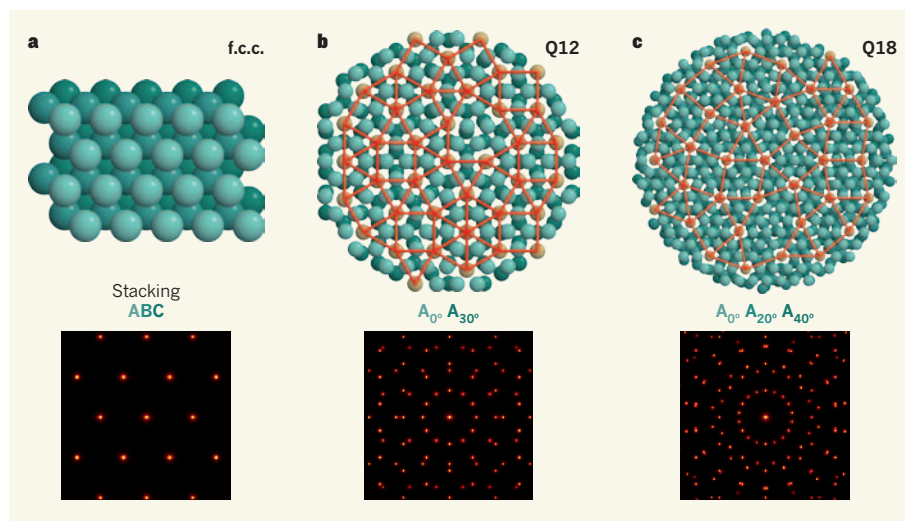


Figure 1 | Periodic and aperiodic micellar phases. **a**, Periodic face-centred cubic crystal phase (f.c.c.). **b, c**, Aperiodic quasicrystalline phases with 12-fold and 18-fold rotational symmetries (Q12, Q18). Each of these three phases can be described as a stacking of hexagonal layers of micelles (turquoise in **a**, and turquoise and yellow in **b** and **c**). In f.c.c., three flat layers A, B, C are shifted. Fischer *et al.*¹ model Q12 as two, on average hexagonal, layers that are rotated by 30° relative to each other (A_0 , A_{30}), and Q18 as three layers that are rotated by 0°, 20° and 40° (A_0 , A_{20} , A_{40}). The layers are not flat, and instead, the yellow micelles are positioned in between layer pairs. Red bonds indicate an exemplary tiling, and simulated diffraction patterns (bottom) indicate the phases' rotational symmetry.

M. ENGEL

symmetries. (The structure of a material with n -fold rotational symmetry remains unchanged under rotation by an angle of $360^\circ/n$ around the symmetry axis.) Other rotational symmetries are possible in quasicrystals. By far the most commonly found quasicrystals in alloys are icosahedral, or axial with decagonal (10-fold) symmetry. Other, more rarely observed quasicrystals are octagonal (8-fold) and dodecagonal (12-fold). Interestingly, until now, all reported soft-matter quasicrystals had 12-fold symmetry. Computer simulations show how dodecagonal symmetries are preferred in micellar systems because of entropy⁸. Indeed, one of the two quasicrystal phases discovered by Fischer and colleagues is 12-fold symmetric (Q12; Fig. 1b).

Much more surprising is their report of an additional phase with 18-fold symmetry (Q18; Fig. 1c) from the analysis of the quasicrystals' X-ray and neutron diffraction patterns, which Fischer *et al.*¹ classified as an enneagonal (9-fold) quasicrystal. Because diffraction patterns are always inversion symmetric, it is an open question whether their quasicrystal has 9-fold or 18-fold symmetry in real space. Although not prohibited, 18-fold (or 9-fold) rotational symmetry has never before been reported.

The preference for certain symmetries may be understood by recognizing that quasicrystals can be embedded into higher-dimensional periodic lattices. The minimal embedding dimension of an n -fold symmetric structure is four for $n = 5, 8, 10$ and 12, and six for $n = 7, 9, 14$ and 18. All known quasicrystals have a low minimal embedding dimension of four. Perhaps, then, quasicrystals with $n = 5, 8, 10$ and 12 are easier to form than other quasicrystals⁹. Indeed, Fischer and colleagues' 18-fold structure is the first report of a quasicrystal with a novel rotational symmetry in more than 20 years.

Proper identification of soft-matter quasicrystals is challenging because they tend to be more thermally disordered than atomic crystals. This has important implications for the interpretation of diffraction patterns. As previously shown¹⁰, the scattering conditions responsible for the appearance of the Bragg diffraction peaks that characterize order in a material are highly sensitive to structural imperfections. These imperfections can lead to the appearance of secondary peaks. Fischer *et al.*¹ find an f.c.c. phase, at higher temperatures than for either Q12 or Q18, that has pronounced secondary peaks and thus many structural defects. The transition to the quasicrystal on cooling involves a change of the layer stacking in f.c.c., which would be signified by a transformation of the secondary peaks into primary peaks. With the authors' currently available diffraction data, however, one cannot definitively confirm this transformation scenario.

Quasicrystal identification from diffraction

experiments also suffers from the fact that if a system contains two orientations of a crystal, then the diffraction pattern is the superposition of the two respective diffraction patterns. In many cases, the constituent crystals can share a very specific orientational relationship. For example, an f.c.c. twin can seem to have 12-fold symmetry. Fischer *et al.* ruled out twinning by comparing the experimental diffraction patterns with simulated diffraction patterns of an f.c.c. twin and a quasicrystal model. In future work, real-space images of the quasicrystals, captured by cryo-scanning electron microscopy and transmission electron microscopy, will allow even more definitive quasicrystal identification. That the observed structures are quasicrystals, and not twinned structures, is further supported by the observed phase transitions: from f.c.c. (at temperatures above 25 °C) to Q12 (at 20–15 °C) to Q18 (at less than 10 °C). Such a sequence is not expected from a system that undergoes twinning.

The discovery of Q18 constitutes the first account of a novel quasicrystalline structure on the nanoscale with no equivalent known for atomic crystals. It has the highest order of diffraction symmetry recorded so far in any single-domain crystal. Such an order of symmetry may be relevant for photonic (quasi-)crystals¹¹. Photonic crystals affect the propagation of light in the same way as the potential in a semiconductor crystal affects electron motion by defining a region of forbidden electronic

energies, the bandgap. In photonics, a high degree of rotational symmetry is desirable for uniform bandgaps. The discovery of this new quasicrystal is even more remarkable — the system is a chemically well known, aqueous solution of simple block copolymers common in industrial applications as wetting, dispersing and foaming agents. The finding should spur a closer look at micelle-forming systems and inspire, in an interesting twist, the search for atomic analogues of soft-matter structures. ■

Sharon C. Glotzer and Michael Engel are in the Department of Chemical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA. S.C.G. is also in the Department of Materials Science and Engineering. e-mail: sglotzer@umich.edu

1. Fischer, S. *et al.* *Proc. Natl Acad. Sci. USA* **108**, 1810–1814 (2011).
2. Zihler, P. & Kamien, R. D. *Phys. Rev. Lett.* **85**, 3528–3531 (2000).
3. Zeng, X. *et al.* *Nature* **428**, 157–160 (2004).
4. Talapin, D. V. *et al.* *Nature* **461**, 964–967 (2009).
5. Haji-Akbari, A. *et al.* *Nature* **462**, 773–777 (2009).
6. Hayashida, K., Dotera, T., Takano, A. & Matsushita, Y. *Phys. Rev. Lett.* **98**, 195502 (2007).
7. Lee, S., Bluemle, M. J. & Bates, F. S. *Science* **330**, 349–353 (2010).
8. Iacovella, C. R., Keys, A. S. & Glotzer, S. C. Preprint at <http://arxiv.org/abs/1102.5589> (2011).
9. Mikhalev, J. *et al.* *Proc. Natl Acad. Sci. USA* **107**, 7214–7218 (2010).
10. Förster, S. *et al.* *Nature Mater.* **6**, 888–893 (2007).
11. Man, W., Megens, M., Steinhardt, P. J. & Chaikin, P. M. *Nature* **436**, 993–996 (2005).

PROGRAMMED CELL DEATH

Apoptosis meets necrosis

Apoptotic cell death is essential for the development of multicellular organisms. Paradoxically, three proteins instrumental in apoptosis also collaborate to preserve life by preventing necrotic cell death. SEE LETTERS P.363, P.368 & P.373

MARCUS E. PETER

Apoptotic cell death can be induced by two distinct pathways: one intrinsic to cells and one extrinsic. Among the main players in the extrinsic pathway in both mice and humans are the adapter protein FADD, the death-executing protease enzyme caspase-8, and a regulator of caspase-8 activity, FLIP. These death-promoting proteins are also involved in embryonic development. How can such apparently opposing functions coexist? Three papers in this issue^{1–3} provide evidence that the proteins act together to suppress another type of programmed cell death — necrosis.

The extrinsic pathway of apoptosis begins with the binding of an appropriate ligand to members of the death-receptor family, which lie in the cell membrane. These then recruit and activate the death-inducing signalling complex (DISC). FADD, caspase-8 and FLIP are all essential components of DISC.

Deletion of the genes encoding each of these DISC proteins causes mice to die *in utero* at mid-gestation as a result of vascular, cardiac and blood-cell-formation defects. Also, *in vitro* proliferation of immune cells called T cells is impaired if any of the three genes is deleted. What's more, tissue-specific deletions of caspase-8 unveiled immunity-related functions of this enzyme: the prevention of

symmetries. (The structure of a material with n -fold rotational symmetry remains unchanged under rotation by an angle of $360^\circ/n$ around the symmetry axis.) Other rotational symmetries are possible in quasicrystals. By far the most commonly found quasicrystals in alloys are icosahedral, or axial with decagonal (10-fold) symmetry. Other, more rarely observed quasicrystals are octagonal (8-fold) and dodecagonal (12-fold). Interestingly, until now, all reported soft-matter quasicrystals had 12-fold symmetry. Computer simulations show how dodecagonal symmetries are preferred in micellar systems because of entropy⁸. Indeed, one of the two quasicrystal phases discovered by Fischer and colleagues is 12-fold symmetric (Q12; Fig. 1b).

Much more surprising is their report of an additional phase with 18-fold symmetry (Q18; Fig. 1c) from the analysis of the quasicrystals' X-ray and neutron diffraction patterns, which Fischer *et al.*¹ classified as an enneagonal (9-fold) quasicrystal. Because diffraction patterns are always inversion symmetric, it is an open question whether their quasicrystal has 9-fold or 18-fold symmetry in real space. Although not prohibited, 18-fold (or 9-fold) rotational symmetry has never before been reported.

The preference for certain symmetries may be understood by recognizing that quasicrystals can be embedded into higher-dimensional periodic lattices. The minimal embedding dimension of an n -fold symmetric structure is four for $n = 5, 8, 10$ and 12 , and six for $n = 7, 9, 14$ and 18 . All known quasicrystals have a low minimal embedding dimension of four. Perhaps, then, quasicrystals with $n = 5, 8, 10$ and 12 are easier to form than other quasicrystals⁹. Indeed, Fischer and colleagues' 18-fold structure is the first report of a quasicrystal with a novel rotational symmetry in more than 20 years.

Proper identification of soft-matter quasicrystals is challenging because they tend to be more thermally disordered than atomic crystals. This has important implications for the interpretation of diffraction patterns. As previously shown¹⁰, the scattering conditions responsible for the appearance of the Bragg diffraction peaks that characterize order in a material are highly sensitive to structural imperfections. These imperfections can lead to the appearance of secondary peaks. Fischer *et al.*¹ find an f.c.c. phase, at higher temperatures than for either Q12 or Q18, that has pronounced secondary peaks and thus many structural defects. The transition to the quasicrystal on cooling involves a change of the layer stacking in f.c.c., which would be signified by a transformation of the secondary peaks into primary peaks. With the authors' currently available diffraction data, however, one cannot definitively confirm this transformation scenario.

Quasicrystal identification from diffraction

experiments also suffers from the fact that if a system contains two orientations of a crystal, then the diffraction pattern is the superposition of the two respective diffraction patterns. In many cases, the constituent crystals can share a very specific orientational relationship. For example, an f.c.c. twin can seem to have 12-fold symmetry. Fischer *et al.* ruled out twinning by comparing the experimental diffraction patterns with simulated diffraction patterns of an f.c.c. twin and a quasicrystal model. In future work, real-space images of the quasicrystals, captured by cryo-scanning electron microscopy and transmission electron microscopy, will allow even more definitive quasicrystal identification. That the observed structures are quasicrystals, and not twinned structures, is further supported by the observed phase transitions: from f.c.c. (at temperatures above 25°C) to Q12 (at 20 – 15°C) to Q18 (at less than 10°C). Such a sequence is not expected from a system that undergoes twinning.

The discovery of Q18 constitutes the first account of a novel quasicrystalline structure on the nanoscale with no equivalent known for atomic crystals. It has the highest order of diffraction symmetry recorded so far in any single-domain crystal. Such an order of symmetry may be relevant for photonic (quasi-)crystals¹¹. Photonic crystals affect the propagation of light in the same way as the potential in a semiconductor crystal affects electron motion by defining a region of forbidden electronic

energies, the bandgap. In photonics, a high degree of rotational symmetry is desirable for uniform bandgaps. The discovery of this new quasicrystal is even more remarkable — the system is a chemically well known, aqueous solution of simple block copolymers common in industrial applications as wetting, dispersing and foaming agents. The finding should spur a closer look at micelle-forming systems and inspire, in an interesting twist, the search for atomic analogues of soft-matter structures. ■

Sharon C. Glotzer and Michael Engel are in the Department of Chemical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA. S.C.G. is also in the Department of Materials Science and Engineering. e-mail: sglotzer@umich.edu

1. Fischer, S. *et al.* *Proc. Natl Acad. Sci. USA* **108**, 1810–1814 (2011).
2. Zihler, P. & Kamien, R. D. *Phys. Rev. Lett.* **85**, 3528–3531 (2000).
3. Zeng, X. *et al.* *Nature* **428**, 157–160 (2004).
4. Talapin, D. V. *et al.* *Nature* **461**, 964–967 (2009).
5. Haji-Akbari, A. *et al.* *Nature* **462**, 773–777 (2009).
6. Hayashida, K., Dotera, T., Takano, A. & Matsushita, Y. *Phys. Rev. Lett.* **98**, 195502 (2007).
7. Lee, S., Bluemle, M. J. & Bates, F. S. *Science* **330**, 349–353 (2010).
8. Iacovella, C. R., Keys, A. S. & Glotzer, S. C. Preprint at <http://arxiv.org/abs/1102.5589> (2011).
9. Mikhalev, J. *et al.* *Proc. Natl Acad. Sci. USA* **107**, 7214–7218 (2010).
10. Förster, S. *et al.* *Nature Mater.* **6**, 888–893 (2007).
11. Man, W., Megens, M., Steinhardt, P. J. & Chaikin, P. M. *Nature* **436**, 993–996 (2005).

PROGRAMMED CELL DEATH

Apoptosis meets necrosis

Apoptotic cell death is essential for the development of multicellular organisms. Paradoxically, three proteins instrumental in apoptosis also collaborate to preserve life by preventing necrotic cell death. SEE LETTERS P.363, P.368 & P.373

MARCUS E. PETER

Apoptotic cell death can be induced by two distinct pathways: one intrinsic to cells and one extrinsic. Among the main players in the extrinsic pathway in both mice and humans are the adapter protein FADD, the death-executing protease enzyme caspase-8, and a regulator of caspase-8 activity, FLIP. These death-promoting proteins are also involved in embryonic development. How can such apparently opposing functions coexist? Three papers in this issue^{1–3} provide evidence that the proteins act together to suppress another type of programmed cell death — necrosis.

The extrinsic pathway of apoptosis begins with the binding of an appropriate ligand to members of the death-receptor family, which lie in the cell membrane. These then recruit and activate the death-inducing signalling complex (DISC). FADD, caspase-8 and FLIP are all essential components of DISC.

Deletion of the genes encoding each of these DISC proteins causes mice to die *in utero* at mid-gestation as a result of vascular, cardiac and blood-cell-formation defects. Also, *in vitro* proliferation of immune cells called T cells is impaired if any of the three genes is deleted. What's more, tissue-specific deletions of caspase-8 unveiled immunity-related functions of this enzyme: the prevention of

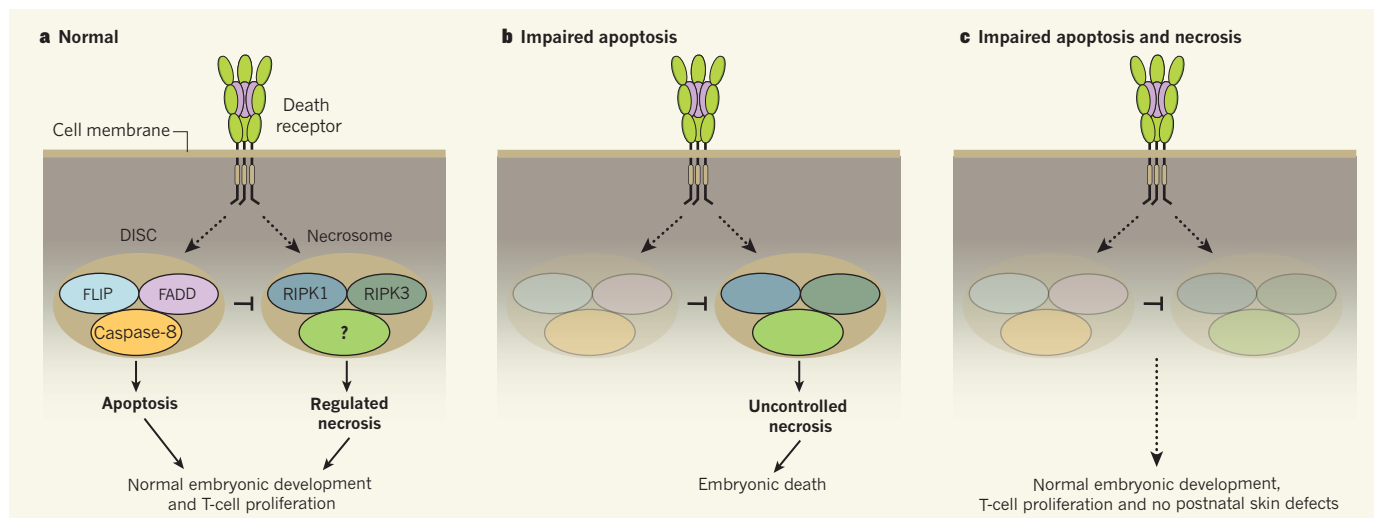


Figure 1 | Programmed-cell-death pathways are intertwined. **a**, In normal mice, death-receptor-mediated apoptosis and necrosis exist in a balance that is in part accomplished by cleavage of the necrosome components RIPK1 and RIPK3 by components of the apoptotic DISC complex. **b**, In mice deficient in any of the DISC proteins FADD, caspase-8 and FLIP, uncontrolled necrosis causes cell death and defects in various tissues, leading to embryonic death.

c, New work^{1–3} shows that, in mice lacking any of the three DISC proteins and either RIPK1 or RIPK3, neither apoptosis nor necrosis seems to occur, preventing the defects described in **b**. Nonetheless, mice deficient in both FADD and RIPK1 die shortly after birth, and those lacking both caspase-8 and RIPK3 develop a lymphoaccumulation disease due to the absence of death-receptor-mediated apoptosis.

skin inflammation, generation of myeloid and lymphoid cells, and differentiation of macrophages. The question remained whether these various defects were due to loss of one or several functions of the DISC components. The new reports^{1–3} suggest that almost all defects seen in any of the DISC-gene mutant mice can be traced back to one activity: the suppression of necrosis.

But how does necrosis relate to apoptosis? Although apoptosis has been synonymous with programmed cell death for many years, there are at least three main forms of programmed cell death: autophagy, necrosis and apoptosis. Autophagy allows a cell to survive — for instance, under conditions of limited nutrient availability — by digesting its own components; if carried to exhaustion, this results in cell death. Necrosis, in contrast to apoptosis, has been viewed as a form of accidental cell death brought about by injury to the cell by pathogens or toxins.

Although several studies have suggested that certain forms of necrosis are programmed, the big breakthrough in the study of programmed necrosis (also called necroptosis⁴) came from the discovery that the enzymatic activity of the protein RIPK1 is instrumental in the execution of necrosis⁵. Intriguingly, RIPK1 was originally cloned through its interaction with CD95, a death receptor that induces apoptosis.

More recently, RIPK3 — another member of the RIPK family — was found^{6,7} to function with RIPK1 in a complex called the necrosome. Caspase-8 can cleave both RIPK1 and RIPK3, inhibiting their role in caspase-independent cell death^{8,9}. This finding suggested crosstalk between DISC and the necrosome.

To generate mice lacking both a DISC and a necrosome component (double-knockout mice), Oberst *et al.* (page 363)¹ and Kaiser

et al. (page 368)² crossed caspase-8-deficient mice with RIPK3-deficient mice; Zhang and colleagues (page 373)³ crossed FADD-deficient mice with RIPK1-deficient mice. Strikingly, in all cases, the embryonic lethality observed in the absence of either FADD or caspase-8 was completely corrected. This observation alone confirms that FADD and caspase-8 on one side, and RIPK1 and RIPK3 on the other, are components of two opposing death mechanisms; the former seem to suppress the latter during embryonic development (Fig. 1).

Oberst *et al.*¹ and Kaiser *et al.*² also report that T-cell proliferation in their double-knockout mice was normal. This suggests that, in the absence of a gene encoding any DISC protein, uncontrolled necrosis inhibits T-cell proliferation. Moreover, none of the double-knockout mice showed chronic skin inflammation, and the development of various immune cells was also normal.

Why are genes of the necrosome expressed, only to be inhibited during embryonic development? After all, without a functional necrosome, mice make it through embryonic development just fine¹⁰. Part of the answer probably lies in an additional activity of RIPK1 — to activate the transcription factor NF- κ B. In Zhang and colleagues' study³, RIPK1-deficient mice died shortly after birth, and this neonatal mortality could not be corrected by additional deletion of FADD. Moreover, Oberst *et al.*¹ and Kaiser *et al.*² find that, with age, mice lacking both RIPK3 and caspase-8 suffered from a disorder of the immune system called progressive lymphoaccumulation, which also occurs in CD95-deficient mice¹¹.

Apoptosis and necrosis seem to be regulated by different forms of caspase-8. The induction

of apoptosis requires a fully processed, heterotetrameric form of caspase-8. However, mice carrying a mutant caspase-8 that prevents release of the mature enzyme, and so renders it inactive for mediating apoptosis, develop normally, showing no defects in T-cell proliferation¹²; this hints at suppression of the necrosome.

Although none of the teams^{1–3} generated mice deficient in both FLIP and RIPK1 or FLIP and RIPK3, Oberst *et al.*¹ tackled this question *in vitro*. They demonstrate that FLIP, independently of its activity in regulating caspase-8, is required for the protective function of the DISC genes during embryonic development. The caspase-8–FLIP complex seems to participate in preventing the formation of a FADD–RIPK1–RIPK3-containing complex in response to TNF — a factor that induces necrosis. These data suggest that mice lacking FLIP and either RIPK1 or RIPK3 also may not show defects in development or T-cell proliferation.

On the basis of these data^{1–3}, it is tempting to conclude that all of the non-apoptotic effects of caspase-8 are due to its role in cleaving the necrosome. Let's not forget, however, that caspase-8 has various non-apoptotic activities, ranging from growth promotion, cell motility and invasiveness, to tumour metastasis. Investigation is needed into whether all of these activities can be traced back to suppression of the necrosome. The simple view that caspase-8 controls the necrosome through proteolytic cleavage of RIPK1 and/or RIPK3 is not fully supported by the available data¹³. And Zhang and colleagues' study suggests that FADD-deficient embryos show greatly increased levels of RIPK1 as well as signs of RIPK3 aggregation, which may not be solely

caused by a lack of cleavage by caspase-8.

Little is known about pathways downstream of the necrosome. Either reactive oxygen species or activation of autophagy through deregulated metabolism might be essential for the completion of programmed necrosis¹⁴. It has become clear, however, that the various death pathways are more closely intertwined than was previously thought. The crosstalk between the components regulating autophagy¹⁵, apoptosis and necrosis^{1–3} amply demonstrate that. ■

Marcus E. Peter is at the *Feinberg School of Medicine, Division of Hematology/Oncology, Northwestern University, Chicago, Illinois 60611, USA.*

e-mail: m-peter@northwestern.edu

1. Oberst, A. *et al. Nature* **471**, 363–367 (2011).
2. Kaiser, W. J. *et al. Nature* **471**, 368–372 (2011).
3. Zhang, J. *et al. Nature* **471**, 373–376 (2011).
4. Degterev, A. *et al. Nature Chem. Biol.* **4**, 313–321 (2008).
5. Holler, N. *et al. Nature Immunol.* **1**, 489–495 (2000).
6. He, S. *et al. Cell* **137**, 1100–1111 (2009).
7. Cho, Y. *et al. Cell* **137**, 1112–1123 (2009).
8. Lin, Y., Devin, A., Rodriguez, Y. & Liu, Z. *Genes Dev.* **13**, 2514–2526 (1999).
9. Feng, S. *et al. Cell Signal.* **19**, 2056–2067 (2007).
10. Newton, K., Sun, X. & Dixit, V. M. *Mol. Cell Biol.* **24**, 1464–1469 (2004).
11. Adachi, M. *et al. Nature Genet.* **11**, 294–300 (1995).
12. Kang, T.-B. *et al. J. Immunol.* **181**, 2522–2532 (2008).
13. Osborn, S. L. *et al. Proc. Natl Acad. Sci. USA* **107**, 13034–13039 (2010).
14. Zhang, D.-W. *et al. Science* **325**, 332–336 (2009).
15. Maiuri, M. C. *et al. Nature Rev. Mol. Cell Biol.* **8**, 741–752 (2007).

readily than other crustal constituents at the high temperatures found at depth³.

The study region is the western Cordillera, reaching from the San Andreas fault system in California to the Rocky Mountains in Wyoming and Colorado. The apparent association of low v_p/v_s with zones of recently active plate-boundary deformation in this region suggests that high abundance of quartz-rich rocks in the lower crust, often described as a weak zone of jelly in the continental sandwich of crust and mantle⁴, is a key ingredient for initiating deformation (Fig. 1). Once a deformation zone has been initiated, further weakening by rising temperatures and addition of fluids may sustain a permanently weak zone over many plate-tectonic cycles. Such inherited weakness may isolate continental interiors from deformation and force the repeated reactivation of plate-boundary faults during alternating cycles of plate divergence and convergence⁵.

The method developed by Lowry and Pérez-Gussinyé relies on an automated data-analysis product⁶, which calculates bulk crustal properties (v_p/v_s and crustal thickness, H) at several hundreds of seismic stations using converted pressure-to-shear-wave arrivals resolved on the radial component of motion. Those parameters are notoriously difficult to estimate accurately because of a well-known trade-off between H and v_p/v_s . To improve the resolution of the estimates, the authors use two additional constraints from statistical inference: optimal spatial interpolation and gravity modelling. The model uses gravity anomalies to obtain an optimal density structure that fits variations in v_p/v_s , H and additional contributions from heat-flow data. The technique thus cleverly combines disconnected data sets with statistical modelling to significantly improve the accuracy of the solution.

Nevertheless, the method is limited by the initial automated solution, which is obtained assuming a single-layer, isotropic crust with a flat crust–mantle boundary. Any deviation from this simple model, produced by non-horizontal structure, anisotropy or multiple layering, may produce complicated patterns of shear-wave arrivals on both radial and transverse components of motion; when inverted for the single-layer crust, interference patterns result that can bias the solution. One possible improvement to the method would be to use transverse-component shear-wave arrivals both to evaluate the validity of the single-layer crustal model and to help refine it.

There is some debate on where the strength of tectonic plates resides. In the traditional jelly-sandwich model of continental plate strength, a weak lower crust underlies a strong and brittle upper crust and overlies a strong uppermost mantle layer⁴. An alternative view is that the uppermost mantle contributes little to continental strength in many regions, leaving the brittle crust alone to support tectonic stresses⁷.

Studies that model geotectonically measured

EARTH SCIENCE

Continental jelly

An approach integrating different data sets has been used to map out seismic-velocity ratios in the crust of western North America. High inferred quartz content correlates with tectonic deformation zones. SEE LETTER P.353

ROLAND BÜRGMANN & PASCAL AUDET

Ever since the recognition of plate tectonics on Earth, with its jigsaw puzzle of shifting plates, geoscientists have wondered why otherwise strong and rigid continents repeatedly break up and collide along the same zones of apparent weakness. On page 353 of this issue, Lowry and Pérez-Gussinyé¹ propose that the inherent weakness of these persistent deformation zones may be caused by the low strength of quartz, and its relative abundance in such zones in the continental crust — Earth's outermost layer, which is generally

30–50 kilometres thick in continental regions.

The authors' data were drawn from the EarthScope Transportable Array of seismic stations in western North America, with additional constraints coming from gravity and heat-flow measurements. They used a new approach to develop a map of the ratio between pressure-wave (v_p) and shear-wave (v_s) propagation velocities in the crust from distant earthquakes (see Fig. 3 on page 355). On the basis of experimental data², Lowry and Pérez-Gussinyé argue that low values of this ratio (v_p/v_s of about 1.8 or lower) are uniquely associated with high concentrations of quartz in the crust, a mineral that flows much more

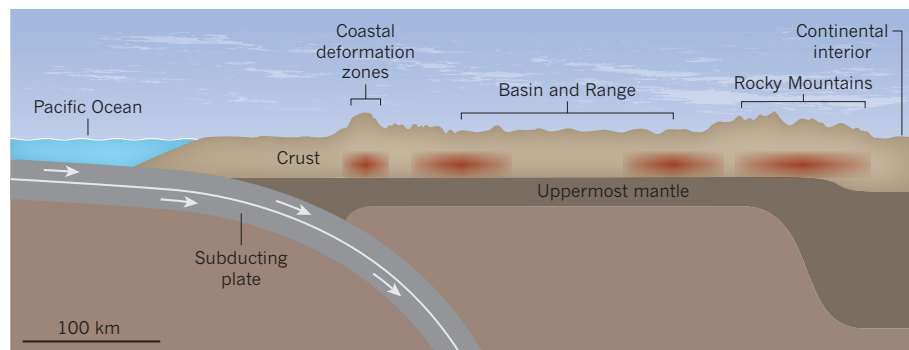


Figure 1 | Cross-section through zones of active deformation in western North America and the stable continental interior. At the left (west) is the subduction zone; at right, the stable continental interior. Lowry and Pérez-Gussinyé¹ find that regions of high inferred quartz content in the lower crust (red) correlate well with zones of recent tectonism exemplified by the coastal deformation zones, parts of the Basin and Range province and the Rocky Mountains. The thickness of the North American plate (crust and rigid uppermost mantle) decreases from more than 200 km in the interior to about 50 km in the west as a result of thermal thinning and hydration caused by past and present subduction.

caused by a lack of cleavage by caspase-8.

Little is known about pathways downstream of the necrosome. Either reactive oxygen species or activation of autophagy through deregulated metabolism might be essential for the completion of programmed necrosis¹⁴. It has become clear, however, that the various death pathways are more closely intertwined than was previously thought. The crosstalk between the components regulating autophagy¹⁵, apoptosis and necrosis^{1–3} amply demonstrate that. ■

Marcus E. Peter is at the *Feinberg School of Medicine, Division of Hematology/Oncology, Northwestern University, Chicago, Illinois 60611, USA.*

e-mail: m-peter@northwestern.edu

1. Oberst, A. *et al. Nature* **471**, 363–367 (2011).
2. Kaiser, W. J. *et al. Nature* **471**, 368–372 (2011).
3. Zhang, J. *et al. Nature* **471**, 373–376 (2011).
4. Degterev, A. *et al. Nature Chem. Biol.* **4**, 313–321 (2008).
5. Holler, N. *et al. Nature Immunol.* **1**, 489–495 (2000).
6. He, S. *et al. Cell* **137**, 1100–1111 (2009).
7. Cho, Y. *et al. Cell* **137**, 1112–1123 (2009).
8. Lin, Y., Devin, A., Rodriguez, Y. & Liu, Z. *Genes Dev.* **13**, 2514–2526 (1999).
9. Feng, S. *et al. Cell Signal.* **19**, 2056–2067 (2007).
10. Newton, K., Sun, X. & Dixit, V. M. *Mol. Cell Biol.* **24**, 1464–1469 (2004).
11. Adachi, M. *et al. Nature Genet.* **11**, 294–300 (1995).
12. Kang, T.-B. *et al. J. Immunol.* **181**, 2522–2532 (2008).
13. Osborn, S. L. *et al. Proc. Natl Acad. Sci. USA* **107**, 13034–13039 (2010).
14. Zhang, D.-W. *et al. Science* **325**, 332–336 (2009).
15. Maiuri, M. C. *et al. Nature Rev. Mol. Cell Biol.* **8**, 741–752 (2007).

readily than other crustal constituents at the high temperatures found at depth³.

The study region is the western Cordillera, reaching from the San Andreas fault system in California to the Rocky Mountains in Wyoming and Colorado. The apparent association of low v_p/v_s with zones of recently active plate-boundary deformation in this region suggests that high abundance of quartz-rich rocks in the lower crust, often described as a weak zone of jelly in the continental sandwich of crust and mantle⁴, is a key ingredient for initiating deformation (Fig. 1). Once a deformation zone has been initiated, further weakening by rising temperatures and addition of fluids may sustain a permanently weak zone over many plate-tectonic cycles. Such inherited weakness may isolate continental interiors from deformation and force the repeated reactivation of plate-boundary faults during alternating cycles of plate divergence and convergence⁵.

The method developed by Lowry and Pérez-Gussinyé relies on an automated data-analysis product⁶, which calculates bulk crustal properties (v_p/v_s and crustal thickness, H) at several hundreds of seismic stations using converted pressure-to-shear-wave arrivals resolved on the radial component of motion. Those parameters are notoriously difficult to estimate accurately because of a well-known trade-off between H and v_p/v_s . To improve the resolution of the estimates, the authors use two additional constraints from statistical inference: optimal spatial interpolation and gravity modelling. The model uses gravity anomalies to obtain an optimal density structure that fits variations in v_p/v_s , H and additional contributions from heat-flow data. The technique thus cleverly combines disconnected data sets with statistical modelling to significantly improve the accuracy of the solution.

Nevertheless, the method is limited by the initial automated solution, which is obtained assuming a single-layer, isotropic crust with a flat crust–mantle boundary. Any deviation from this simple model, produced by non-horizontal structure, anisotropy or multiple layering, may produce complicated patterns of shear-wave arrivals on both radial and transverse components of motion; when inverted for the single-layer crust, interference patterns result that can bias the solution. One possible improvement to the method would be to use transverse-component shear-wave arrivals both to evaluate the validity of the single-layer crustal model and to help refine it.

There is some debate on where the strength of tectonic plates resides. In the traditional jelly-sandwich model of continental plate strength, a weak lower crust underlies a strong and brittle upper crust and overlies a strong uppermost mantle layer⁴. An alternative view is that the uppermost mantle contributes little to continental strength in many regions, leaving the brittle crust alone to support tectonic stresses⁷.

Studies that model geotectonically measured

EARTH SCIENCE

Continental jelly

An approach integrating different data sets has been used to map out seismic-velocity ratios in the crust of western North America. High inferred quartz content correlates with tectonic deformation zones. SEE LETTER P.353

ROLAND BÜRGMANN & PASCAL AUDET

Ever since the recognition of plate tectonics on Earth, with its jigsaw puzzle of shifting plates, geoscientists have wondered why otherwise strong and rigid continents repeatedly break up and collide along the same zones of apparent weakness. On page 353 of this issue, Lowry and Pérez-Gussinyé¹ propose that the inherent weakness of these persistent deformation zones may be caused by the low strength of quartz, and its relative abundance in such zones in the continental crust — Earth's outermost layer, which is generally

30–50 kilometres thick in continental regions.

The authors' data were drawn from the EarthScope Transportable Array of seismic stations in western North America, with additional constraints coming from gravity and heat-flow measurements. They used a new approach to develop a map of the ratio between pressure-wave (v_p) and shear-wave (v_s) propagation velocities in the crust from distant earthquakes (see Fig. 3 on page 355). On the basis of experimental data², Lowry and Pérez-Gussinyé argue that low values of this ratio (v_p/v_s of about 1.8 or lower) are uniquely associated with high concentrations of quartz in the crust, a mineral that flows much more

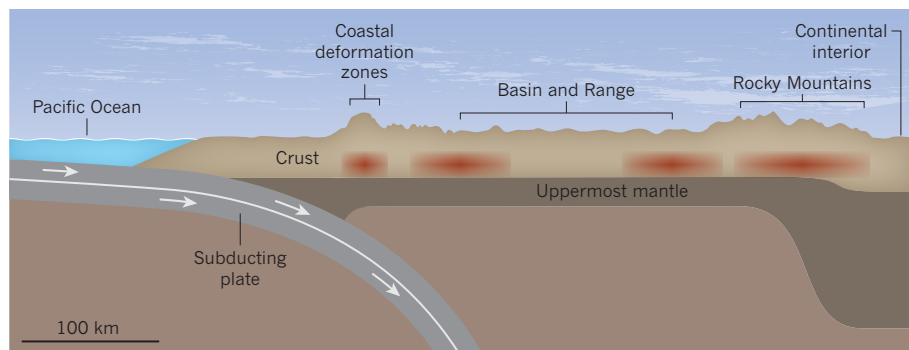


Figure 1 | Cross-section through zones of active deformation in western North America and the stable continental interior. At the left (west) is the subduction zone; at right, the stable continental interior. Lowry and Pérez-Gussinyé¹ find that regions of high inferred quartz content in the lower crust (red) correlate well with zones of recent tectonism exemplified by the coastal deformation zones, parts of the Basin and Range province and the Rocky Mountains. The thickness of the North American plate (crust and rigid uppermost mantle) decreases from more than 200 km in the interior to about 50 km in the west as a result of thermal thinning and hydration caused by past and present subduction.

deformation due to viscous flow of rocks induced by large earthquakes in western North America indicate that, at these short timescales, the lower crust is quite strong, whereas the upper mantle below about 50 km is much weaker^{3,8}. The effective flow strengths of the lower crust found in the geodetic studies seem too high for the quartz-rich make-up suggested by Lowry and Pérez-Gussinyé. On the other hand, low, long-term static strength, inferred from estimates of effective elastic plate thickness of only 10 km or less (see Fig. 4 on p. 356), is consistent with a weak lower crust in the region. Thus, lower-crustal viscosities at very long (millions of years) timescales may effectively control the stability of continental crust and upper mantle⁸.

Is a quartz-rich layer in the crust, only tens of kilometres thick, able to initiate break-up of a continental plate originally dominated by a strong mantle layer up to 200 km thick⁹? Lowry and Pérez-Gussinyé argue that, following initial deformation enabled by the quartz-rich crust, the strong mantle layer can be further softened by high temperatures and/or fluids derived from subducting oceanic plates¹⁰, leading to the eventual loss or soggy of the bottom slice of the jelly sandwich. Importantly, even where the uppermost mantle remains stable, as indicated by high estimates of elastic-plate thickness, a quartz-weakened lower crust can promote tectonic deformation. This may have been the case during the most recent tectonic period of the northern Rocky Mountains (Fig. 1), the Laramide orogeny¹.

The validity of the model can be tested when similar techniques are applied to different tectonic provinces that have experienced successive cycles of supercontinent formation and mountain-building. In particular, as Lowry and Pérez-Gussinyé suggest, the extension of the Transportable Array across older orogens in eastern North America during the next two years will provide a unique opportunity to test the role of quartz in mountain building. ■

Roland Bürgmann and Pascal Audet are in the Department of Earth and Planetary Science, and the Berkeley Seismological Laboratory, University of California, Berkeley, California 94720-4767, USA.
e-mail: burgmann@seismo.berkeley.edu

- Lowry, A. R. & Pérez-Gussinyé, M. *Nature* **471**, 353–359 (2011).
- Christensen, N. I. *J. Geophys. Res.* **101**, 3139–3156 (1996).
- Bürgmann, R. & Dresen, G. *Annu. Rev. Earth Planet. Sci.* **36**, 531–567 (2008).
- Chen, W. P. & Molnar, P. *J. Geophys. Res.* **88**, 4183–4214 (1983).
- Audet, P. & Bürgmann, R. *Nature Geosci.* **4**, 184–187 (2011).
- www.iris.edu/dms/products/ears
- Jackson, J. *GSA Today* **12**, 4–10 (2002).
- Thatcher, W. & Pollitz, F. F. *GSA Today* **18**, 4–11 (2008).
- Yuan, H. & Romanowicz, B. *Nature* **466**, 1063–1068 (2009).
- Dixon, J. E., Dixon, T. H., Bell, D. R. & Malservisi, R. *Earth Planet. Sci. Lett.* **222**, 451–467 (2004).

REPRODUCTIVE BIOLOGY

Progesterone's gateway into sperm

The hormone progesterone rapidly activates intracellular signalling in human sperm, regulating key aspects of their physiology. An ion channel unique to the sperm tail seems to relay progesterone's signal. SEE LETTERS P.382 & P.387

STEVE PUBLICOVER
& CHRISTOPHER BARRATT

The ovarian hormone progesterone classically binds to a nuclear receptor, initiating gene transcription. But how does it stimulate the transcriptionally inactive human spermatozoon in preparation for fertilization? This question has long both fascinated and frustrated reproductive biologists. In this issue, Strünker and colleagues¹ and Lishko *et al.*² provide an unexpected answer: progesterone activates a sperm-specific calcium ion (Ca^{2+}) channel called CatSper.

For a sperm to reach the egg, it must penetrate the cumulus oophorus, a thick layer around the egg composed of granulosa cells embedded in a gelatinous matrix. These cells actively synthesize progesterone, such that its concentration within the cumulus is in the micromolar range. It was first reported more than 20 years ago³ that progesterone, even at concentrations well below those present in the cumulus, induces immediate influx into human sperm of Ca^{2+} — a factor central to regulation of sperm function^{4,5}. Progesterone is therefore believed to have a crucial role during the events leading to fertilization⁶.

Sperm cells respond to progesterone within less than a second, which is characteristic of classical signalling pathways that involve cell-surface receptors^{3,6}. Such non-nuclear actions of steroid hormones are quite common. In fact, progesterone and its related hormones are considered to have two distinct modes of action: through intracellular nuclear receptors, which regulate transcription; and through non-genomic receptors, probably at the plasma membrane, which regulate ion channels, G-protein-coupled receptors and signalling pathways mediated by kinase enzymes⁷. However, the mechanism of progesterone-induced Ca^{2+} influx in sperm has resisted all attempts at characterization, with even the type of 'receptor', let alone the nature of the Ca^{2+} -influx pathway, remaining a mystery. This has been particularly frustrating because the phenomenon is probably of considerable clinical significance: in human sperm, failure of progesterone-activated Ca^{2+} influx is correlated with reduced fertility⁶.

The solution to this mystery follows directly

from two crucial advances in the field. First, in 2001 two groups^{8,9} reported the discovery of the Ca^{2+} -permeable cation channel (CatSper), which is expressed only in the plasma membrane of a domain in the sperm tail called the principal piece. Sperm from genetically manipulated mice that cannot express CatSper have impaired motility and, crucially, cannot display hyperactivation — an extravagant, highly asymmetric form of flagellar beating that is regulated by Ca^{2+} and is essential for fertilization. CatSper-deficient male mice are infertile.

The second, more recent, advance was the development of a method for applying to sperm the technique of whole-cell patch clamping, which records ionic currents across the entire plasma membrane of a cell. Using this technique, researchers showed that increased alkalinity of the sperm cytoplasm strongly activates CatSper channels, promoting Ca^{2+} flux into the cell. Strünker *et al.* (page 382) and Lishko and colleagues (page 387) now use this powerful technique to elucidate the mechanism by which progesterone induces rapid Ca^{2+} influx into human sperm.

Progesterone-induced membrane currents have identical characteristics to those carried by CatSper. For instance, the biophysical aspects of the currents are indistinguishable, with both progesterone and increased intracellular pH stimulating CatSper by shifting its voltage sensitivity so that it opens at lower voltages (Fig. 1). What's more, pharmacological manipulation has the same effects both on CatSper currents activated by increasing intracellular pH and on those stimulated by progesterone; applied together, progesterone and increased alkalinity act synergistically^{1,2}.

The effect of progesterone on CatSper is not simply a nonspecific effect of steroid hormones: another steroid hormone, oestradiol, has no effect on this channel². However, several prostaglandins — non-protein mediators that increase intracellular Ca^{2+} concentration in human sperm — have strikingly similar effects to progesterone. Moreover, Strünker and colleagues' measurements of intracellular Ca^{2+} concentration in progesterone-stimulated sperm showed that compounds that block CatSper currents also reduce the progesterone-induced rise in Ca^{2+} concentration, and that

deformation due to viscous flow of rocks induced by large earthquakes in western North America indicate that, at these short timescales, the lower crust is quite strong, whereas the upper mantle below about 50 km is much weaker^{3,8}. The effective flow strengths of the lower crust found in the geodetic studies seem too high for the quartz-rich make-up suggested by Lowry and Pérez-Gussinyé. On the other hand, low, long-term static strength, inferred from estimates of effective elastic plate thickness of only 10 km or less (see Fig. 4 on p. 356), is consistent with a weak lower crust in the region. Thus, lower-crustal viscosities at very long (millions of years) timescales may effectively control the stability of continental crust and upper mantle⁸.

Is a quartz-rich layer in the crust, only tens of kilometres thick, able to initiate break-up of a continental plate originally dominated by a strong mantle layer up to 200 km thick⁹? Lowry and Pérez-Gussinyé argue that, following initial deformation enabled by the quartz-rich crust, the strong mantle layer can be further softened by high temperatures and/or fluids derived from subducting oceanic plates¹⁰, leading to the eventual loss or soggy of the bottom slice of the jelly sandwich. Importantly, even where the uppermost mantle remains stable, as indicated by high estimates of elastic-plate thickness, a quartz-weakened lower crust can promote tectonic deformation. This may have been the case during the most recent tectonic period of the northern Rocky Mountains (Fig. 1), the Laramide orogeny¹.

The validity of the model can be tested when similar techniques are applied to different tectonic provinces that have experienced successive cycles of supercontinent formation and mountain-building. In particular, as Lowry and Pérez-Gussinyé suggest, the extension of the Transportable Array across older orogens in eastern North America during the next two years will provide a unique opportunity to test the role of quartz in mountain building. ■

Roland Bürgmann and Pascal Audet are in the Department of Earth and Planetary Science, and the Berkeley Seismological Laboratory, University of California, Berkeley, California 94720-4767, USA.
e-mail: burgmann@seismo.berkeley.edu

- Lowry, A. R. & Pérez-Gussinyé, M. *Nature* **471**, 353–359 (2011).
- Christensen, N. I. *J. Geophys. Res.* **101**, 3139–3156 (1996).
- Bürgmann, R. & Dresen, G. *Annu. Rev. Earth Planet. Sci.* **36**, 531–567 (2008).
- Chen, W. P. & Molnar, P. *J. Geophys. Res.* **88**, 4183–4214 (1983).
- Audet, P. & Bürgmann, R. *Nature Geosci.* **4**, 184–187 (2011).
- www.iris.edu/dms/products/ears
- Jackson, J. *GSA Today* **12**, 4–10 (2002).
- Thatcher, W. & Pollitz, F. F. *GSA Today* **18**, 4–11 (2008).
- Yuan, H. & Romanowicz, B. *Nature* **466**, 1063–1068 (2009).
- Dixon, J. E., Dixon, T. H., Bell, D. R. & Malservisi, R. *Earth Planet. Sci. Lett.* **222**, 451–467 (2004).

REPRODUCTIVE BIOLOGY

Progesterone's gateway into sperm

The hormone progesterone rapidly activates intracellular signalling in human sperm, regulating key aspects of their physiology. An ion channel unique to the sperm tail seems to relay progesterone's signal. SEE LETTERS P.382 & P.387

STEVE PUBLICOVER
& CHRISTOPHER BARRATT

The ovarian hormone progesterone classically binds to a nuclear receptor, initiating gene transcription. But how does it stimulate the transcriptionally inactive human spermatozoon in preparation for fertilization? This question has long both fascinated and frustrated reproductive biologists. In this issue, Strücker and colleagues¹ and Lishko *et al.*² provide an unexpected answer: progesterone activates a sperm-specific calcium ion (Ca^{2+}) channel called CatSper.

For a sperm to reach the egg, it must penetrate the cumulus oophorus, a thick layer around the egg composed of granulosa cells embedded in a gelatinous matrix. These cells actively synthesize progesterone, such that its concentration within the cumulus is in the micromolar range. It was first reported more than 20 years ago³ that progesterone, even at concentrations well below those present in the cumulus, induces immediate influx into human sperm of Ca^{2+} — a factor central to regulation of sperm function^{4,5}. Progesterone is therefore believed to have a crucial role during the events leading to fertilization⁶.

Sperm cells respond to progesterone within less than a second, which is characteristic of classical signalling pathways that involve cell-surface receptors^{3,6}. Such non-nuclear actions of steroid hormones are quite common. In fact, progesterone and its related hormones are considered to have two distinct modes of action: through intracellular nuclear receptors, which regulate transcription; and through non-genomic receptors, probably at the plasma membrane, which regulate ion channels, G-protein-coupled receptors and signalling pathways mediated by kinase enzymes⁷. However, the mechanism of progesterone-induced Ca^{2+} influx in sperm has resisted all attempts at characterization, with even the type of 'receptor', let alone the nature of the Ca^{2+} -influx pathway, remaining a mystery. This has been particularly frustrating because the phenomenon is probably of considerable clinical significance: in human sperm, failure of progesterone-activated Ca^{2+} influx is correlated with reduced fertility⁶.

The solution to this mystery follows directly

from two crucial advances in the field. First, in 2001 two groups^{8,9} reported the discovery of the Ca^{2+} -permeable cation channel (CatSper), which is expressed only in the plasma membrane of a domain in the sperm tail called the principal piece. Sperm from genetically manipulated mice that cannot express CatSper have impaired motility and, crucially, cannot display hyperactivation — an extravagant, highly asymmetric form of flagellar beating that is regulated by Ca^{2+} and is essential for fertilization. CatSper-deficient male mice are infertile.

The second, more recent, advance was the development of a method for applying to sperm the technique of whole-cell patch clamping, which records ionic currents across the entire plasma membrane of a cell. Using this technique, researchers showed that increased alkalinity of the sperm cytoplasm strongly activates CatSper channels, promoting Ca^{2+} flux into the cell. Strücker *et al.* (page 382) and Lishko and colleagues (page 387) now use this powerful technique to elucidate the mechanism by which progesterone induces rapid Ca^{2+} influx into human sperm.

Progesterone-induced membrane currents have identical characteristics to those carried by CatSper. For instance, the biophysical aspects of the currents are indistinguishable, with both progesterone and increased intracellular pH stimulating CatSper by shifting its voltage sensitivity so that it opens at lower voltages (Fig. 1). What's more, pharmacological manipulation has the same effects both on CatSper currents activated by increasing intracellular pH and on those stimulated by progesterone; applied together, progesterone and increased alkalinity act synergistically^{1,2}.

The effect of progesterone on CatSper is not simply a nonspecific effect of steroid hormones: another steroid hormone, oestradiol, has no effect on this channel². However, several prostaglandins — non-protein mediators that increase intracellular Ca^{2+} concentration in human sperm — have strikingly similar effects to progesterone. Moreover, Strücker and colleagues' measurements of intracellular Ca^{2+} concentration in progesterone-stimulated sperm showed that compounds that block CatSper currents also reduce the progesterone-induced rise in Ca^{2+} concentration, and that

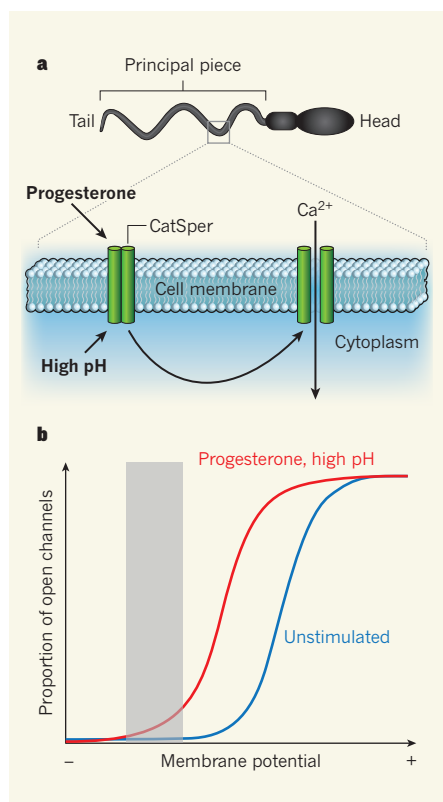


Figure 1 | CatSper and calcium-ion influx.

a, The CatSper ion channel, which occurs in the plasma membrane of the principal-piece domain of the sperm tail, allows Ca^{2+} influx into these cells. New work^{1,2} shows that progesterone leads to the opening of this channel, probably either by binding to it directly or through an associated protein; increased intracellular pH has the same effect. On opening, CatSper channels allow Ca^{2+} entry into the cell, which regulates events vital for fertilization. **b**, Opening of CatSper channels depends on the electrical difference across the cell membrane (the membrane potential) and occurs when the cell becomes electrically more positive inside. The normal membrane potential in sperm (grey bar) is such that nearly all CatSper channels are closed (blue line). Both progesterone and increased intracellular pH 'shift' the electrical sensitivity of CatSper so that the channel can open at more negative membrane potentials (red trace).

stringent buffering of this ion in the external medium abolishes the response to progesterone¹. Intriguingly, the efficacy of progesterone is increased by *in vitro* manipulations designed to induce sperm capacitation (a crucial maturation process that naturally occurs in the female reproductive tract before fertilization).

The non-genomic action of progesterone is much more potent in human sperm than in mouse sperm². But why? Lishko and colleagues show that, at the intracellular pH of 7.0 (a value within the physiological range), mouse spermatozoa show notable CatSper currents, whereas human sperm show a much smaller current. On applying progesterone to human sperm, the current increases to a level closely resembling that in mouse sperm, but in mouse

sperm stimulation with this hormone leads to no further increase in current². It seems, therefore, that in human sperm, progesterone induces a modulation of CatSper function that in mouse sperm is constitutive (at least under the conditions used in these experiments). This is potentially a crucial species difference in sperm regulation within the female reproductive tract.

The two papers also present a much clearer idea of how progesterone exerts its effect by modulating CatSper. Lishko *et al.*² could record progesterone-induced currents even in isolated sperm tails, which precludes indirect effects of progesterone exerted through receptors on the sperm head. Furthermore, Strücker *et al.*¹ provide compelling evidence that progesterone does not stimulate synthesis of the signalling molecule cyclic AMP, and they couldn't detect any effects of manipulating cAMP levels on Ca^{2+} influx through the sperm membrane. These observations rule out involvement of the cAMP–protein kinase A signalling cascade in the progesterone–CatSper response.

The new data also suggest that progesterone directly activates CatSper, by binding either to the channel itself or to an associated subunit(s). Whether CatSper activation is the only effect of progesterone on Ca^{2+} -signalling in human sperm remains to be seen. Several previous studies have attempted to identify progesterone receptors⁶. Both novel receptors and truncated versions of the classical (nuclear-type) receptors (some of these apparently localized to the sperm head) were proposed to mediate the effects of this hormone. Although such

receptors almost certainly do not contribute to the modulation of CatSper reported here, it is noteworthy that completely blocking CatSper currents inhibits — but does not abolish — the effect of progesterone on intracellular Ca^{2+} levels^{1,10}.

Mobilization of intracellular Ca^{2+} stores, leading to complex Ca^{2+} signalling, occurs in progesterone-stimulated human sperm⁵. Is this purely a downstream effect of CatSper activation or does progesterone activate a separate pathway? Are store-controlled Ca^{2+} channels involved? These two studies^{1,2} provide exciting insights, and there is more to come. ■

Steve Publicover is in the School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK.

Christopher Barratt is in the School of Medicine, University of Dundee, Ninewells Hospital, Dundee DD1 9SY, UK.

e-mails: s.j.publicover@bham.ac.uk; c.barratt@dundee.ac.uk

1. Strücker, T. *et al.* *Nature* **471**, 382–386 (2011).
2. Lishko, P. V., Botchkina, I. L. & Kirichok, Y. *Nature* **471**, 387–391 (2011).
3. Blackmore, P. F., Beebe, S. J., Danforth, D. R. & Alexander, N. J. *Biol. Chem.* **265**, 1376–1380 (1990).
4. Darszon, A. *et al.* *Int. Rev. Cytol.* **243**, 79–172 (2005).
5. Publicover, S., Harper, C. V. & Barratt, C. *Nature Cell Biol.* **9**, 235–242 (2007).
6. Baldi, E. *et al.* *Mol. Cell Endocrinol.* **308**, 39–46 (2009).
7. Simoncini, T. & Genazzani, A. R. *Eur. J. Endocrinol.* **148**, 281–292 (2003).
8. Quill, T. A., Ren, D., Clapham, D. E. & Garbers, D. L. *Proc. Natl Acad. Sci. USA* **98**, 12527–12531 (2001).
9. Ren, D. *et al.* *Nature* **413**, 603–609 (2001).
10. Garcia, M. A. & Meizel, S. *Biol. Reprod.* **60**, 102–109 (1999).

HIGH-TEMPERATURE SUPERCONDUCTIVITY

The secret of the hourglass

The finding that a cobalt oxide insulator's magnetism is similar to that of cuprate superconductors lends support to the popular but contentious idea that stripe-like electronic order is present in the latter materials. SEE LETTER P.341

JAN ZAAEN

One hundred years after its discovery, superconductivity is still an active field of research. On page 341 of this issue, Boothroyd *et al.*¹ describe experimental results on an insulating material that offer insight into the physics of one of the most intriguing families of superconductors — the copper oxides, or cuprates.

Conventional superconductivity — that which occurs in simple metals such as lead and aluminium — was explained back in

1957 by Bardeen, Cooper and Schrieffer, in what is known as the BCS theory². But in 1986, a different, high-temperature form of superconductivity was discovered in complex cuprates³. This discovery rumbled like an earthquake through the physics community, because the superconducting transition temperatures (T_c), below which these materials conduct electricity without resistance, were much too high to be explained by BCS theory. What causes superconductivity in the cuprates is still much of a mystery, but intensive research has shown that the ground rules of

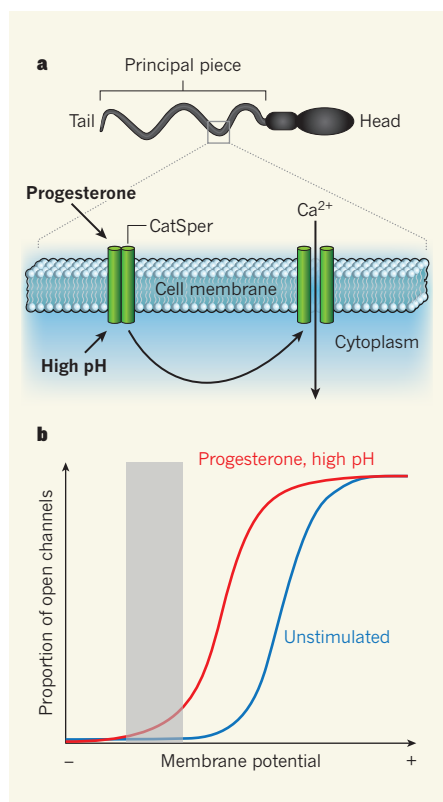


Figure 1 | CatSper and calcium-ion influx.

a, The CatSper ion channel, which occurs in the plasma membrane of the principal-piece domain of the sperm tail, allows Ca^{2+} influx into these cells. New work^{1,2} shows that progesterone leads to the opening of this channel, probably either by binding to it directly or through an associated protein; increased intracellular pH has the same effect. On opening, CatSper channels allow Ca^{2+} entry into the cell, which regulates events vital for fertilization. **b**, Opening of CatSper channels depends on the electrical difference across the cell membrane (the membrane potential) and occurs when the cell becomes electrically more positive inside. The normal membrane potential in sperm (grey bar) is such that nearly all CatSper channels are closed (blue line). Both progesterone and increased intracellular pH 'shift' the electrical sensitivity of CatSper so that the channel can open at more negative membrane potentials (red trace).

stringent buffering of this ion in the external medium abolishes the response to progesterone¹. Intriguingly, the efficacy of progesterone is increased by *in vitro* manipulations designed to induce sperm capacitation (a crucial maturation process that naturally occurs in the female reproductive tract before fertilization).

The non-genomic action of progesterone is much more potent in human sperm than in mouse sperm². But why? Lishko and colleagues show that, at the intracellular pH of 7.0 (a value within the physiological range), mouse spermatozoa show notable CatSper currents, whereas human sperm show a much smaller current. On applying progesterone to human sperm, the current increases to a level closely resembling that in mouse sperm, but in mouse

sperm stimulation with this hormone leads to no further increase in current². It seems, therefore, that in human sperm, progesterone induces a modulation of CatSper function that in mouse sperm is constitutive (at least under the conditions used in these experiments). This is potentially a crucial species difference in sperm regulation within the female reproductive tract.

The two papers also present a much clearer idea of how progesterone exerts its effect by modulating CatSper. Lishko *et al.*² could record progesterone-induced currents even in isolated sperm tails, which precludes indirect effects of progesterone exerted through receptors on the sperm head. Furthermore, Strünker *et al.*¹ provide compelling evidence that progesterone does not stimulate synthesis of the signalling molecule cyclic AMP, and they couldn't detect any effects of manipulating cAMP levels on Ca^{2+} influx through the sperm membrane. These observations rule out involvement of the cAMP–protein kinase A signalling cascade in the progesterone–CatSper response.

The new data also suggest that progesterone directly activates CatSper, by binding either to the channel itself or to an associated subunit(s). Whether CatSper activation is the only effect of progesterone on Ca^{2+} -signalling in human sperm remains to be seen. Several previous studies have attempted to identify progesterone receptors⁶. Both novel receptors and truncated versions of the classical (nuclear-type) receptors (some of these apparently localized to the sperm head) were proposed to mediate the effects of this hormone. Although such

receptors almost certainly do not contribute to the modulation of CatSper reported here, it is noteworthy that completely blocking CatSper currents inhibits — but does not abolish — the effect of progesterone on intracellular Ca^{2+} levels^{1,10}.

Mobilization of intracellular Ca^{2+} stores, leading to complex Ca^{2+} signalling, occurs in progesterone-stimulated human sperm⁵. Is this purely a downstream effect of CatSper activation or does progesterone activate a separate pathway? Are store-controlled Ca^{2+} channels involved? These two studies^{1,2} provide exciting insights, and there is more to come. ■

Steve Publicover is in the School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK.

Christopher Barratt is in the School of Medicine, University of Dundee, Ninewells Hospital, Dundee DD1 9SY, UK.

e-mails: s.j.publicover@bham.ac.uk; c.barratt@dundee.ac.uk

1. Strünker, T. *et al.* *Nature* **471**, 382–386 (2011).
2. Lishko, P. V., Botchkina, I. L. & Kirichok, Y. *Nature* **471**, 387–391 (2011).
3. Blackmore, P. F., Beebe, S. J., Danforth, D. R. & Alexander, N. J. *Biol. Chem.* **265**, 1376–1380 (1990).
4. Darszon, A. *et al.* *Int. Rev. Cytol.* **243**, 79–172 (2005).
5. Publicover, S., Harper, C. V. & Barratt, C. *Nature Cell Biol.* **9**, 235–242 (2007).
6. Baldi, E. *et al.* *Mol. Cell Endocrinol.* **308**, 39–46 (2009).
7. Simoncini, T. & Genazzani, A. R. *Eur. J. Endocrinol.* **148**, 281–292 (2003).
8. Quill, T. A., Ren, D., Clapham, D. E. & Garbers, D. L. *Proc. Natl Acad. Sci. USA* **98**, 12527–12531 (2001).
9. Ren, D. *et al.* *Nature* **413**, 603–609 (2001).
10. Garcia, M. A. & Meizel, S. *Biol. Reprod.* **60**, 102–109 (1999).

HIGH-TEMPERATURE SUPERCONDUCTIVITY

The secret of the hourglass

The finding that a cobalt oxide insulator's magnetism is similar to that of cuprate superconductors lends support to the popular but contentious idea that stripe-like electronic order is present in the latter materials. SEE LETTER P.341

JAN ZANEN

One hundred years after its discovery, superconductivity is still an active field of research. On page 341 of this issue, Boothroyd *et al.*¹ describe experimental results on an insulating material that offer insight into the physics of one of the most intriguing families of superconductors — the copper oxides, or cuprates.

Conventional superconductivity — that which occurs in simple metals such as lead and aluminium — was explained back in

1957 by Bardeen, Cooper and Schrieffer, in what is known as the BCS theory². But in 1986, a different, high-temperature form of superconductivity was discovered in complex cuprates³. This discovery rumbled like an earthquake through the physics community, because the superconducting transition temperatures (T_c), below which these materials conduct electricity without resistance, were much too high to be explained by BCS theory. What causes superconductivity in the cuprates is still much of a mystery, but intensive research has shown that the ground rules of

the quantum physics governing the electron systems in the cuprates are very different from those of conventional systems: in contrast to the featureless quantum gas formed by the electrons in conventional superconductors, the high- T_c electrons seem to form highly organized types of ‘quantum matter’, exhibiting a richness of electronic behaviour that challenges even the diversity found in classical complex fluids⁴.

With their unique ability to measure how the magnetic properties of this ‘electron stuff’ fluctuate in time, experiments based on the scattering of neutrons have played a key part in discerning some of the signatures of the cuprates’ electron quantum matter. Such measurements revealed that these magnetic fluctuations behave in a peculiarly organized manner, nicknamed the hourglass spectrum (Fig. 1). Boothroyd *et al.*¹ now demonstrate that the magnetic fluctuations in an insulating material, for which it is established that the electrons freeze out in a complex ‘stripe’ pattern, look very similar to those of the cuprates, thereby supporting the popular but controversial idea that such stripes are also present in superconducting cuprates — although in a quantum ‘dynamical’ form⁵.

In conventional superconductors, electronic quantum zero-point motions dominate to such a degree that the electrons form a nearly featureless quantum gas, just leaving room for the formation of the electron pairs that, according to BCS theory, are responsible for superconductivity². Such pairs are also formed in the cuprates, albeit in a mysteriously sturdy form. But there is much more going on in these materials. A large body of experimental work has shown that the cuprates’ high- T_c superconducting state seems to coexist³, to a greater or lesser degree, with a wealth of exotic quantum organizational phenomena, including static and dynamical stripes, spontaneous ‘diamagnetic’ currents, and quantum nematics (quantum versions of the liquid crystals used in liquid-crystal displays).

Among these exotic forms of electronic order, the stripes have a special status because they were the first to be identified — in fact as a surprising output⁶ of computer simulations that I generated as a young postdoc in 1987. This story starts with the particulars of Cu ions⁴, having the effect that the electrons in the copper oxide lattice repel each other much more strongly than in conventional superconductors. As a result, the cuprate electron world is more like traffic on a congested highway than like the ‘quantum fog’ of conventional metals. In the stoichiometric cuprates, the electron-traffic density is so high that the traffic jams completely, forming what is known as a Mott insulator. To get the electrons moving, some of them are removed by chemical doping, and one enters an ‘underdoped’ regime of quantum mechanical stop-and-go traffic — and here the exotic orders emerge.

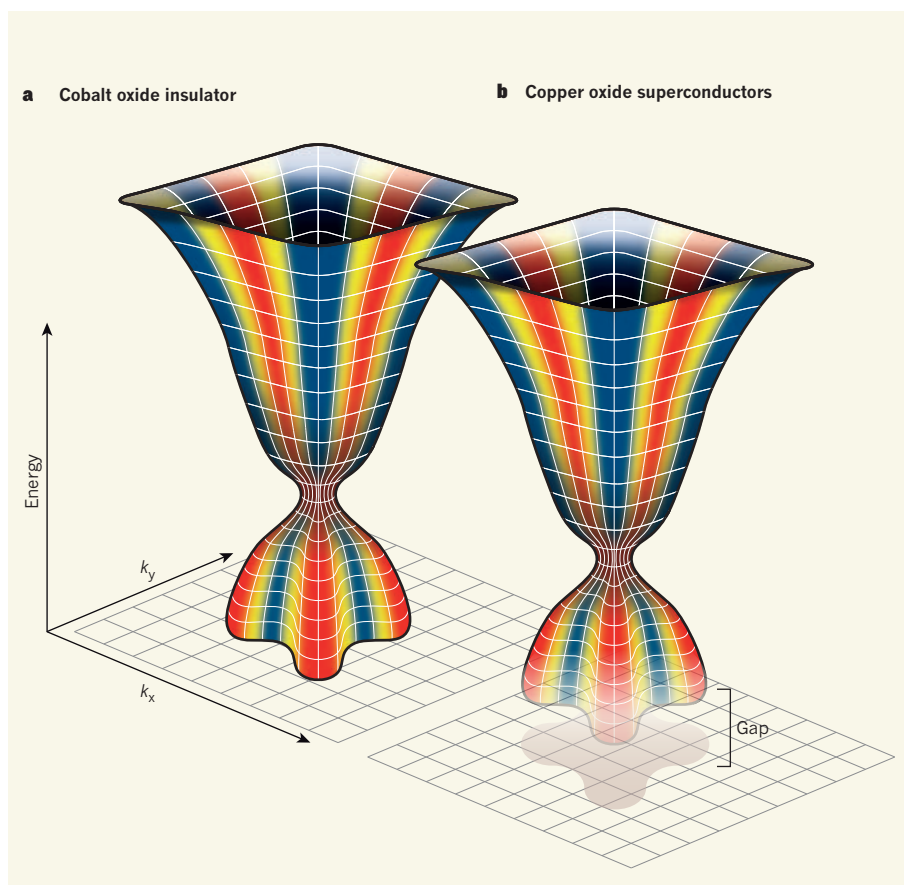


Figure 1 | The hourglass magnetic spectrum. **a**, Boothroyd *et al.*¹ find that the collective magnetic ‘vibrations’ (spin waves) in a cobalt oxide insulator have an hourglass shape when plotted as a function of their energy and inverse wavelengths (k_x , k_y); the colours indicate the intensity of the spin waves as measured from neutron scattering. **b**, In the superconducting copper oxides, the neutron-scattering results display a very similar shape, except that at low energies a gap appears that is indicative of a collective quantum melting of stripe-like electronic order⁵. (Graphic based on a sketch by A. Boothroyd.)

By dismissing all collective electronic quantum fluctuations, the physics of such dense quantum traffic can be computed, and my calculations⁶ showed that the electron traffic spontaneously forms complex patterns, which we nicknamed stripes. These consist of ‘rivers’, in which the electron motions are relatively free, separated by domains in which the traffic continues to be completely jammed. The electron spins in this stripe phase interact through short-range quantum fluctuations, which cause the spins to freeze out in a special ‘incommensurate’ antiferromagnetic order: within the insulating domains, the orientation of the spin of each electron is opposite to those of its neighbours, and the rivers act as ‘domain walls’ in this antiferromagnet. Disappointingly, in the cuprates, my computations insisted that this static stripe phase should insulate instead of superconduct. But this turned out to have been a blessing in disguise when it became gradually clear in the 1990s that the static stripes explain why doped Mott insulating oxides that do not contain copper as a rule form insulators: invariably, their electrons freeze in the sturdy stripes of my computer code.

Starting from the mid-1990s, evidence accumulated that stripes might also have a role in the cuprates, albeit in a ‘dynamical’ way that is not that well understood: in underdoped superconducting cuprates, stripes would be present in the form of strong correlations at short timescales, but at longer timescales they would fall prey to a quantum melting driven by collective quantum fluctuations^{4,5}. Consistent with this hypothesis, it was found that stripes become static only in some special cases⁷, and that static stripes have a detrimental effect on superconductivity. Neutron-scattering experiments were pivotal in collecting that evidence. These experiments measure spin fluctuations in the materials as a function of their energy and of their wave-number (inverse wavelength), and in this energy–wavenumber space the spin-fluctuation spectrum of the underdoped cuprates is shaped like an hourglass (Fig. 1). It was argued⁵ that this could be explained in terms of collective ‘vibrations’ (spin waves) of the stripe-phase incommensurate antiferromagnetic order, except that at low energies a gap appears in this spectrum that was interpreted as the signature of the collective quantum

melting of the stripe phase as a whole.

However, an equally credible case was made⁸ that the hourglass spectrum could instead be explained in terms of spin excitations in a rather weakly interacting gas of itinerant electrons, and a debate regarding the interpretation of the hourglass spectrum evolved that rages up to the present day. All along, the problem for the dynamical-stripe interpretation was that the modelling of the spin waves involved a lot of assumptions. In this regard, Boothroyd and colleagues' study¹ makes a big difference. The authors perform a neutron-scattering experiment on a material that falls outside the family of cuprate superconductors — a cobalt oxide insulator — and that is known to display stripes⁹ in a simple static form⁶. They show that the material exhibits an hourglass spin-fluctuation spectrum (Fig. 1a) strikingly

similar to that of the cuprates (Fig. 1b); the only difference is seen at low energies, where the cuprate 'quantum gap' is absent in the cobalt oxide. This similarity lends support to the hypothesis that the hourglass spin-fluctuation spectrum in the cuprate superconductors arises from dynamical stripes^{4,5}.

Boothroyd and colleagues' results arrive at a time when the reality of complex quantum matter in underdoped cuprates is becoming mainstream wisdom. Perhaps we already know so much about these materials that research should be refocused on the greatest mystery of all⁴: that increased levels of doping make the complex quantum stuff gradually fade away, and that the best superconductors are found at the point where the electron traffic starts to resemble the quantum fog of the simple metals. ■

Jan Zaanen is at the *Instituut-Lorentz for Theoretical Physics, Leiden University, 2300 RA Leiden, the Netherlands.*
e-mail: jan@lorentz.leidenuniv.nl

1. Boothroyd, A. T., Babkevich, P., Prabhakaran, D. & Freeman, P. G. *Nature* **471**, 341–344 (2011).
2. Schrieffer, J. R. *Theory of Superconductivity* (Perseus, 1999).
3. Bednorz, J. G. & Müller, K. A. *Z. Phys. B* **64**, 189–193 (1986).
4. Zaanen, J. in *100 Years of Superconductivity* (eds Rochalla, H. & Kes, P. H.) (Chapman & Hall, in the press); <http://arxiv.org/abs/1012.5461> (2011).
5. Kivelson, S. A. *et al. Rev. Mod. Phys.* **75**, 1201–1241 (2003).
6. Zaanen, J. & Gunnarsson, O. *Phys. Rev. B* **40**, 7391–7394 (1989).
7. Tranquada, J. M. *et al. Nature* **429**, 534–538 (1995).
8. Eschrig, M. *Adv. Phys.* **55**, 47–183 (2006).
9. Cwik, M. *et al. Phys. Rev. Lett.* **102**, 057201 (2009).

TRANSLATIONAL MEDICINE

Cancer lessons from mice to humans

New clinical trials report the efficacy of two mechanism-based therapies for treating human pancreatic neuroendocrine tumours. Studies in mouse models have contributed to these success stories, and continue to do so.

DAVID TUVESON & DOUGLAS HANAHAN

Advances in cancer medicine have reset our clinical and social expectations: the aim now is to effectively combat formidable tumours — an effort that was previously deemed improbable. Writing in *The New England Journal of Medicine*, Raymond *et al.*¹ and Yao *et al.*² report phase III clinical trials of two drugs that target distinctive cancer-associated signalling pathways. The results suggest an impressive efficacy of both drugs (sunitinib and everolimus) for treating pancreatic neuroendocrine tumours. It is therefore likely that these drugs, which are already standard treatments for other cancers, will become the first new approvals in 25 years by the US Food and Drug Administration (FDA) for treating these cancers, a remarkable milestone.

Pancreatic neuroendocrine tumours (PNET) are uncommon, but difficult to diagnose and treat. These cancers, which originate from the hormone-producing pancreatic islet cells, stand in stark contrast to another type of pancreatic cancer, pancreatic ductal adenocarcinoma, which is much more prevalent and deadly: a larger proportion of patients with PNET undergo surgical excision, and the clinical course of the disease is highly variable. Nonetheless, patients with advanced PNET

who are not candidates for surgery have a terminal illness, and their tumours are difficult to manage; the FDA-approved chemotherapeutic agent streptozotocin shows only modest activity in these patients.

A vast number of potential anticancer drugs are currently in the pipelines of biopharmaceutical companies. Indeed, the scope of mechanism-based targeting is broad, often with several potential drugs affecting the same target. Consequently, it is challenging to decide which targets and candidate drugs might be of value in particular forms of human cancer, especially those that are rare but deadly like PNET. There is growing optimism that genetically engineered mouse models, which can mimic the progression of specific types of human cancer at the genomic and tissue levels, can contribute to this prioritization³. The hope is that preclinical trials of candidate drugs in representative mouse models could help to motivate and guide clinical trials of targeted therapies in the related human tumours (Fig. 1). The two new papers^{1,2} reflect proof of this concept.

The mouse model of PNET, called RIP-Tag2, shows similar tissue-level features to the human tumours⁴. However, the cancer in the animal does not follow the same — currently obscure⁵ — initiating events that lead to human PNET; it is instead driven by a

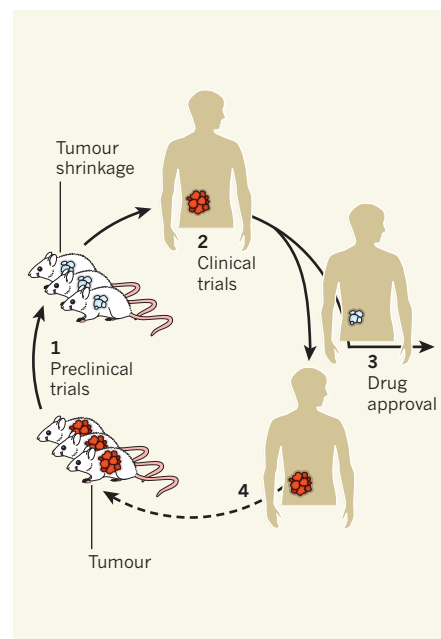


Figure 1 | Linking preclinical and clinical trials. (1) Preclinical trials on cohorts of mice engineered to develop a particular type of cancer are a good starting point for evaluating mechanism-based drugs. (2) If the mice show detectable therapeutic benefits, such as increased survival and/or tumour shrinkage, the preclinical trials can motivate and guide the design of clinical trials in the same type of cancer. (3) Clinical benefit, such as tumour shrinkage, increased progression-free-survival and overall survival, can justify drug approval for clinical use. (4) Relapses and clinical failures, however, can be translated back into refined preclinical trials aiming to understand and circumvent the limitation.

viral oncogene that abrogates the function of two generic tumour-suppressor pathways commonly lost in human tumours.

Preclinical trials in this model had predicted that both sunitinib, a pan-specific inhibitor of tyrosine-kinase enzymes, and everolimus,

melting of the stripe phase as a whole.

However, an equally credible case was made⁸ that the hourglass spectrum could instead be explained in terms of spin excitations in a rather weakly interacting gas of itinerant electrons, and a debate regarding the interpretation of the hourglass spectrum evolved that rages up to the present day. All along, the problem for the dynamical-stripe interpretation was that the modelling of the spin waves involved a lot of assumptions. In this regard, Boothroyd and colleagues' study¹ makes a big difference. The authors perform a neutron-scattering experiment on a material that falls outside the family of cuprate superconductors — a cobalt oxide insulator — and that is known to display stripes⁹ in a simple static form⁶. They show that the material exhibits an hourglass spin-fluctuation spectrum (Fig. 1a) strikingly

similar to that of the cuprates (Fig. 1b); the only difference is seen at low energies, where the cuprate 'quantum gap' is absent in the cobalt oxide. This similarity lends support to the hypothesis that the hourglass spin-fluctuation spectrum in the cuprate superconductors arises from dynamical stripes^{4,5}.

Boothroyd and colleagues' results arrive at a time when the reality of complex quantum matter in underdoped cuprates is becoming mainstream wisdom. Perhaps we already know so much about these materials that research should be refocused on the greatest mystery of all⁴: that increased levels of doping make the complex quantum stuff gradually fade away, and that the best superconductors are found at the point where the electron traffic starts to resemble the quantum fog of the simple metals. ■

Jan Zaanen is at the *Instituut-Lorentz for Theoretical Physics, Leiden University, 2300 RA Leiden, the Netherlands.*
e-mail: jan@lorentz.leidenuniv.nl

1. Boothroyd, A. T., Babkevich, P., Prabhakaran, D. & Freeman, P. G. *Nature* **471**, 341–344 (2011).
2. Schrieffer, J. R. *Theory of Superconductivity* (Perseus, 1999).
3. Bednorz, J. G. & Müller, K. A. *Z. Phys. B* **64**, 189–193 (1986).
4. Zaanen, J. in *100 Years of Superconductivity* (eds Rochalla, H. & Kes, P. H.) (Chapman & Hall, in the press); <http://arxiv.org/abs/1012.5461> (2011).
5. Kivelson, S. A. *et al. Rev. Mod. Phys.* **75**, 1201–1241 (2003).
6. Zaanen, J. & Gunnarsson, O. *Phys. Rev. B* **40**, 7391–7394 (1989).
7. Tranquada, J. M. *et al. Nature* **429**, 534–538 (1995).
8. Eschrig, M. *Adv. Phys.* **55**, 47–183 (2006).
9. Cwik, M. *et al. Phys. Rev. Lett.* **102**, 057201 (2009).

TRANSLATIONAL MEDICINE

Cancer lessons from mice to humans

New clinical trials report the efficacy of two mechanism-based therapies for treating human pancreatic neuroendocrine tumours. Studies in mouse models have contributed to these success stories, and continue to do so.

DAVID TUVESON & DOUGLAS HANAHAN

Advances in cancer medicine have reset our clinical and social expectations: the aim now is to effectively combat formidable tumours — an effort that was previously deemed improbable. Writing in *The New England Journal of Medicine*, Raymond *et al.*¹ and Yao *et al.*² report phase III clinical trials of two drugs that target distinctive cancer-associated signalling pathways. The results suggest an impressive efficacy of both drugs (sunitinib and everolimus) for treating pancreatic neuroendocrine tumours. It is therefore likely that these drugs, which are already standard treatments for other cancers, will become the first new approvals in 25 years by the US Food and Drug Administration (FDA) for treating these cancers, a remarkable milestone.

Pancreatic neuroendocrine tumours (PNET) are uncommon, but difficult to diagnose and treat. These cancers, which originate from the hormone-producing pancreatic islet cells, stand in stark contrast to another type of pancreatic cancer, pancreatic ductal adenocarcinoma, which is much more prevalent and deadly: a larger proportion of patients with PNET undergo surgical excision, and the clinical course of the disease is highly variable. Nonetheless, patients with advanced PNET

who are not candidates for surgery have a terminal illness, and their tumours are difficult to manage; the FDA-approved chemotherapeutic agent streptozotocin shows only modest activity in these patients.

A vast number of potential anticancer drugs are currently in the pipelines of biopharmaceutical companies. Indeed, the scope of mechanism-based targeting is broad, often with several potential drugs affecting the same target. Consequently, it is challenging to decide which targets and candidate drugs might be of value in particular forms of human cancer, especially those that are rare but deadly like PNET. There is growing optimism that genetically engineered mouse models, which can mimic the progression of specific types of human cancer at the genomic and tissue levels, can contribute to this prioritization³. The hope is that preclinical trials of candidate drugs in representative mouse models could help to motivate and guide clinical trials of targeted therapies in the related human tumours (Fig. 1). The two new papers^{1,2} reflect proof of this concept.

The mouse model of PNET, called RIP-Tag2, shows similar tissue-level features to the human tumours⁴. However, the cancer in the animal does not follow the same — currently obscure⁵ — initiating events that lead to human PNET; it is instead driven by a

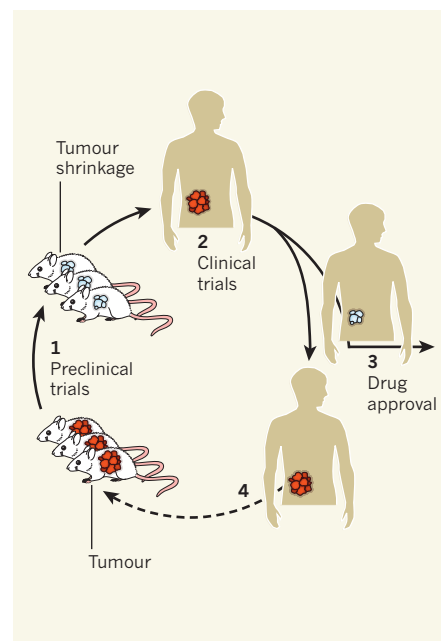


Figure 1 | Linking preclinical and clinical trials. (1) Preclinical trials on cohorts of mice engineered to develop a particular type of cancer are a good starting point for evaluating mechanism-based drugs. (2) If the mice show detectable therapeutic benefits, such as increased survival and/or tumour shrinkage, the preclinical trials can motivate and guide the design of clinical trials in the same type of cancer. (3) Clinical benefit, such as tumour shrinkage, increased progression-free-survival and overall survival, can justify drug approval for clinical use. (4) Relapses and clinical failures, however, can be translated back into refined preclinical trials aiming to understand and circumvent the limitation.

viral oncogene that abrogates the function of two generic tumour-suppressor pathways commonly lost in human tumours.

Preclinical trials in this model had predicted that both sunitinib, a pan-specific inhibitor of tyrosine-kinase enzymes, and everolimus,

which inhibits another kinase, mTOR, would be effective in treating human PNET. Several studies^{6–8} showed that sunitinib, and other kinase inhibitors that target signalling associated with angiogenesis through receptors for the growth factors VEGF and PDGF (thus inhibiting angiogenesis), cause tumour shrinkage. Sunitinib also produced increased survival in the animal studies. These results motivated Raymond and colleagues to perform two phase II trials^{9,10} and now the phase III trial¹ of sunitinib in patients with PNET (Box 1). Similarly, a separate study¹¹ reported the efficacy of another mTOR inhibitor, rapamycin, in treating PNET in the mouse model, presaging the clinical success of everolimus, a refined mTOR inhibitor, which Yao *et al.*² now describe.

Although drug efficacies seen in the pre-clinical trials were encouraging, the trials also revealed limitations — in tumour shrinkage and long-term survival of the mice — that may well influence how these drugs are most effectively used to treat human PNET. Yao and colleagues also find that, whereas everolimus delays time to progression of the disease (progression-free survival), it seemingly does not increase overall survival rates. This trial is still ongoing, however, so the lack of effect on overall survival is not yet conclusive.

Pertinent to this clinical observation is an intriguing result with translational potential from preclinical trials of rapamycin in the mouse model of PNET. Rapamycin on its own produced only a modest overall survival benefit, and the animals also showed evidence of rapamycin resistance following treatment, in the form of regrowth of the previously responding tumours⁹. But when rapamycin was given in combination with another approved drug — erlotinib, which inhibits the growth factor receptor EGFR — the animals' overall survival rate improved significantly and there was a decrease in relapse during treatment¹¹.

These outcomes in the PNET mouse model are consistent with the possibly limited overall survival of patients with PNET following treatment with everolimus only². The preclinical results therefore encourage clinical trials on everolimus in combination with erlotinib (or with other drugs that target downstream effectors in the same signalling pathway). A small clinical trial¹² combining the two drugs to treat PNET is already under way.

For sunitinib, the tumour shrinkage and increased overall survival seen in preclinical trials⁸ are recapitulated in the human trial: Raymond *et al.*¹ report improved both progression-free survival and overall survival after administration of this drug. But, as with everolimus, preclinical trials revealed limitations to the effectiveness of sunitinib in the form of adaptive resistance in PNET. In other words, faced with sunitinib's potent

BOX 1

A closer look at the trials

Both trials^{1,2} enrolled patients with advanced pancreatic neuroendocrine tumours (PNET) who had already received much treatment (including surgery and chemotherapy).

Raymond *et al.*¹ compared the response of 86 randomly selected patients given sunitinib with that of 85 control patients on a placebo drug. Neither the patients nor their doctors were aware of who was receiving which treatment — a double-blind trial.

The median tumour-progression-free survival of patients on sunitinib (11.4 months) was appreciably longer than that of the control group (5.5 months). Moreover, patients treated with sunitinib showed early signs of an increase in overall survival. The side effects were typical of previous experience with sunitinib and included gastrointestinal disturbances and fatigue.

Yao and colleagues² compared the response of 207 randomly selected patients on everolimus with that of 203 patients who received only the best supportive care. In this study, the patients and their physicians were aware of the treatment used, with patients

who were receiving only supportive care having the option to switch to everolimus if their cancer progressed. Nonetheless, all patients were classified by the treatment they initially received (intention to treat).

Everolimus was beneficial, extending the median time of progression from 4.6 months to 11 months. The incidence of rash, gastrointestinal disturbances, fatigue, anaemia and infections was higher in patients treated with this drug, but these side effects were largely manageable. Although no increase in overall survival of patients treated with everolimus was noted, this conclusion is tentative because of the cross-over of patients and the continuance of the trial for the best responders to everolimus.

It should be determined whether these treatments are also beneficial to other groups of patients with PNET, including those who have just had surgery, those who have early-stage disease and those who were marginally too ill to qualify for these trials^{1,2}. **D.T. & D.H.**

effect in blocking angiogenesis, the tumours not only adapt after a period of shrinkage, but also survive the treatment better, inducing alternative pro-angiogenic signalling circuits¹³ and becoming more invasive and metastatic⁸; this reflects a phenomenon seen in other preclinical models as well as in clinical trials^{14–16}.

The mouse data therefore predict eventual failure of therapy with sunitinib alone, and should motivate preclinical and clinical trials to circumvent the evasive resistance — an iterative and bidirectional process of translational therapeutic oncology.

The clinical results with everolimus and sunitinib^{1,2} are landmarks for treating PNET. The approach that led to this — aligned preclinical trials in a representative mouse model and human clinical trials — could also be used to test the efficacy of other anticancer drugs and may well replicate this success story. Indeed, this approach heralds a future in which preclinical trials in genetically engineered mouse models, and in other representative animal models, could guide the development of more effective therapies for human cancers, revealing efficacy, beneficial drug combinations and (potentially surmountable) mechanisms of resistance. ■

David Tuveson is at the Cancer Research UK/ Cambridge Research Institute, University of Cambridge, Cambridge CB2 0RE, UK.

Douglas Hanahan is at the Swiss Institute for Experimental Cancer Research, Swiss Federal Institute of Technology Lausanne, CH-1015 Lausanne, Switzerland.

e-mails: david.tuveson@cancer.org.uk; dh@epfl.ch

1. Raymond, E. *et al.* *N. Engl. J. Med.* **364**, 501–513 (2011).
2. Yao, J. C. *et al.* *N. Engl. J. Med.* **364**, 514–523 (2011).
3. Frese, K. K. & Tuveson, D. A. *Nature Rev. Cancer* **7**, 645–658 (2007).
4. Hanahan, D. *Nature* **315**, 115–122 (1985).
5. Jiao, Y. *et al.* *Science* doi: 10.1126/science.1200609 (2011).
6. Bergers, G., Song, S., Meyer-Morse, N., Bergsland, E. & Hanahan, D. *J. Clin. Invest.* **111**, 1287–1295 (2003).
7. Pietras, K. & Hanahan, D. *J. Clin. Oncol.* **23**, 939–952 (2005).
8. Páez-Ribes, M. *et al.* *Cancer Cell* **15**, 220–231 (2009).
9. Faivre, S. *et al.* *J. Clin. Oncol.* **24**, 25–35 (2006).
10. Kulke, M. H. *et al.* *J. Clin. Oncol.* **26**, 3403–3410 (2008).
11. Chiu, C., Nozawa, H. & Hanahan, D. *J. Clin. Oncol.* **28**, 4425–4433 (2010).
12. <http://clinicaltrials.gov/ct2/show/NCT00843531>
13. Casanovas, O., Hicklin, D. J., Bergers, G. & Hanahan, D. *Cancer Cell* **8**, 299–309 (2005).
14. Bergers, G. & Hanahan, D. *Nature Rev. Cancer* **8**, 592–603 (2008).
15. Azam, F., Mehta, S. & Harris, A. L. *Eur. J. Cancer* **46**, 1323–1332 (2010).
16. Ebos, J. M. L. & Kerbel, R. S. *Nature Rev. Clin. Oncol.* doi:10.1038/nrclinonc.2011.21 (2011).

D.H. declares competing financial interests. See online article for details.

Single-spin addressing in an atomic Mott insulator

Christof Weitenberg¹, Manuel Endres¹, Jacob F. Sherson^{1†}, Marc Cheneau¹, Peter Schauß¹, Takeshi Fukuhara¹, Immanuel Bloch^{1,2} & Stefan Kuhr¹

Ultracold atoms in optical lattices provide a versatile tool with which to investigate fundamental properties of quantum many-body systems. In particular, the high degree of control of experimental parameters has allowed the study of many interesting phenomena, such as quantum phase transitions and quantum spin dynamics. Here we demonstrate how such control can be implemented at the most fundamental level of a single spin at a specific site of an optical lattice. Using a tightly focused laser beam together with a microwave field, we were able to flip the spin of individual atoms in a Mott insulator with sub-diffraction-limited resolution, well below the lattice spacing. The Mott insulator provided us with a large two-dimensional array of perfectly arranged atoms, in which we created arbitrary spin patterns by sequentially addressing selected lattice sites after freezing out the atom distribution. We directly monitored the tunnelling quantum dynamics of single atoms in the lattice prepared along a single line, and observed that our addressing scheme leaves the atoms in the motional ground state. The results should enable studies of entropy transport and the quantum dynamics of spin impurities, the implementation of novel cooling schemes, and the engineering of quantum many-body phases and various quantum information processing applications.

The ability to observe and control the position of single atoms on a surface of a solid by means of scanning tunnelling and atomic force microscopy has revolutionized the field of condensed matter physics^{1,2}. In few-atom systems, coherent control of single particles in, for example, an ion chain, has proven crucial for the implementation of high-fidelity quantum gates and the readout of individual qubits in quantum information processing³. Bringing such levels of control to the regime of large-scale many-body systems has been a long-standing goal in quantum physics. In the context of ultracold atoms in optical lattices, a major challenge has been to combine degenerate atomic samples with single-site addressing resolution and single-atom sensitivity. This full control is essential for many applications in condensed matter physics, such as the study of spin impurities⁴ and quantum spin dynamics^{5,6} within quantum magnetism, entropy transport, the implementation of novel cooling schemes^{7,8} or digital quantum simulations based on Rydberg atoms⁹. For scalable quantum information processing, a Mott insulator with unity filling provides a natural quantum register with several hundreds of qubits. In order to exploit the full potential of such a large-scale system for quantum computation, coherent manipulation of individual spins is indispensable, both within a circuit-based¹⁰ or a one-way quantum computer architecture^{11,12}.

The quest to address atoms on single sites of an optical lattice has a long history^{7,13–22}. In one dimension, single-site addressing has been accomplished optically in a long-wavelength lattice²³, in which however tunnelling was completely suppressed, and also by using magnetic resonance techniques in a sparsely filled short-wavelength lattice^{24,25}. In two dimensions, an electron beam has been used to depopulate sites of a Bose–Einstein condensate loaded into an array of potential tubes, each containing up to 80 atoms²⁶. In this case, coherent spin manipulation was not possible and the readout was done by averaging over more than 100 single images. None of the experiments so far have shown single-atom spin control in strongly correlated systems together with high fidelity single-atom detection.

Here we report the achievement of this goal, by demonstrating single-site-resolved addressing and control of the spin states of individual atoms in a Mott insulator in an optical lattice. The Mott insulator provided us with an almost perfect initial two-dimensional array of atoms in the same spin state. Apart from a few thermal defects, each lattice site contained a single atom in its motional ground state^{27,28}. Using a tightly focused laser beam, we introduced a controlled differential energy shift between two atomic spin states at a given lattice site. Microwave radiation resonant with this shifted transition then allowed us to selectively address the spin of a single atom^{7,17} with high fidelity. We thereby obtained sub-diffraction-limited spatial resolution well below the lattice spacing. By moving the addressing laser beam to different lattice sites and by inducing spin-flips in the Mott insulator, we were able to deterministically create arbitrary two-dimensional spin patterns of individual atoms, thereby realizing a scalable single-atom source^{29–31}. Furthermore, we investigated how much our single-spin manipulation affects the motional state of the atoms by directly monitoring the tunnelling dynamics of single atoms after addressing them. Averaging over several snapshots after different tunnelling times, we fully reconstructed the characteristic spatial probability distribution of the single-atom wavefunction and its coherent evolution over more than 20 lattice sites. We were able to discriminate the dynamics of the atoms in the zeroth and in the first band and found that most of the atoms remained in the motional ground state after addressing.

Experimental set-up

In our experiments, we prepared a two-dimensional sample of ultracold ⁸⁷Rb atoms in an optical lattice, confined in a single antinode of a vertical standing wave along the *z* direction. Two pairs of counterpropagating laser beams (wavelength $\lambda = 1,064$ nm) along the horizontal *x* and *y* directions provided a square lattice with period of $a_{\text{lat}} = \lambda/2 = 532$ nm (for details, see ref. 28). Starting from a Bose–Einstein condensate, we

¹Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany. ²Ludwig-Maximilians-Universität, Schellingstr. 4/II, 80799 München, Germany. [†]Present address: Department of Physics and Astronomy, University of Aarhus, DK-8000 Aarhus C, Denmark.

raised the potential in the x and y lattice axes within 75 ms to values of $V_{x,y} = 23(2) E_r$ (the number in parenthesis denotes the 1σ uncertainty of the last digit), where $E_r = \hbar^2/(2m\lambda^2)$ is the recoil energy, and m denotes the atomic mass of ^{87}Rb . In this way, the interaction energy of the particles with respect to their kinetic energy was increased such that the system undergoes a quantum phase transition to a Mott insulating state^{32–34}. Owing to the external harmonic confinement, the Mott insulator exhibits a shell structure with fixed integer atom numbers increasing in a step-like manner from the outer regions of the system to the inner core^{35,36}. The initial state for all experiments presented here was a single shell with only one atom per lattice site, which in our system was realized for atom numbers smaller than ~ 400 .

We detected the atoms using fluorescence imaging via a high-resolution microscope objective with numerical aperture of $\text{NA} = 0.68$. An optical molasses induced fluorescence light and simultaneously laser-cooled the atoms. Light-assisted collisions lead to rapid losses of atom pairs, such that we only detected the atom number modulo two (that is, either zero or one atom) on each lattice site^{27,28}. With about 5,000 collected photons per atom, we identified individual atoms in the lattice with an excellent signal-to-noise ratio. Even in the regions of high atomic density, we determined the presence or absence of an atom for each lattice site with $>99.5\%$ fidelity using a special reconstruction algorithm²⁸.

Addressing single lattice sites

In order to address the atoms in the lattice, we used an off-resonant laser beam focused by the high-resolution imaging system onto individual lattice sites (Fig. 1). The laser beam causes a differential light shift of the two relevant hyperfine levels $|0\rangle \equiv |F=1, m_F=-1\rangle$ and $|1\rangle \equiv |F=2, m_F=-2\rangle$ and tunes the addressed atom into resonance with an external microwave field at ~ 6.8 GHz. The σ^- -polarized addressing beam had a wavelength of 787.55 nm, between the D_1 and D_2 lines of ^{87}Rb , in order to obtain a large differential light shift between the two hyperfine levels. For perfect σ^- -polarization, this ‘magic’ wavelength generates a light shift only for state $|1\rangle$, and leaves state $|0\rangle$ unaffected. The beam had a diameter of ~ 600 nm full-width at half-maximum (FWHM) and could be moved in the object plane over the entire field of view by changing its angle of incidence into the microscope objective with a two-axis piezo mirror. We were able to position the beam with an accuracy better than $0.1 a_{\text{lat}}$ using an independent calibration measurement of its position together with a feedback that tracks the slowly varying lattice phases (Methods).

If the addressing laser beam is perfectly centred onto one lattice site (Fig. 1b), the differential light shift is $\Delta_{\text{LS}}/(2\pi) \approx -70$ kHz, whereas a neighbouring atom only experiences 10% of the peak intensity. The resulting difference in light shifts can be well resolved spectrally by our microwave (MW) pulses. In order to flip the spin, we performed Landau-Zener sweeps (Methods) of $\sigma_{\text{MW}}/(2\pi) = 60$ kHz width and 20 ms duration, yielding a near flat-top frequency spectrum with a maximum population transfer efficiency of $\sim 95\%$.

As a first experiment, we sequentially flipped the spin of the atoms at selected lattice sites in our Mott insulator with unity filling and spin state $|0\rangle$ (Fig. 2a). The lattice depths were first changed to $V_x = 56 E_r$, $V_y = 90 E_r$ and $V_z = 70 E_r$ in order to completely suppress tunnelling even when the addressing beam locally perturbs the lattice potential. For each lattice site, we then switched on the addressing laser beam with an ‘S’-shaped ramp within 2.5 ms, which is adiabatic with respect to the on-site oscillation frequency of ~ 30 kHz. Subsequently, a microwave pulse with the parameters described above produced spin-flips from $|0\rangle$ to $|1\rangle$. The addressing laser was switched off again within 2.5 ms, before its position was changed in 5 ms to address the next lattice site. For the image of Fig. 2a, this procedure was repeated 16 times in order to flip the spins along a line. Finally, a 5 ms ‘push-out’ laser pulse, resonant with the $F=2$ to $F'=3$ transition, removed all addressed atoms in state $|1\rangle$.

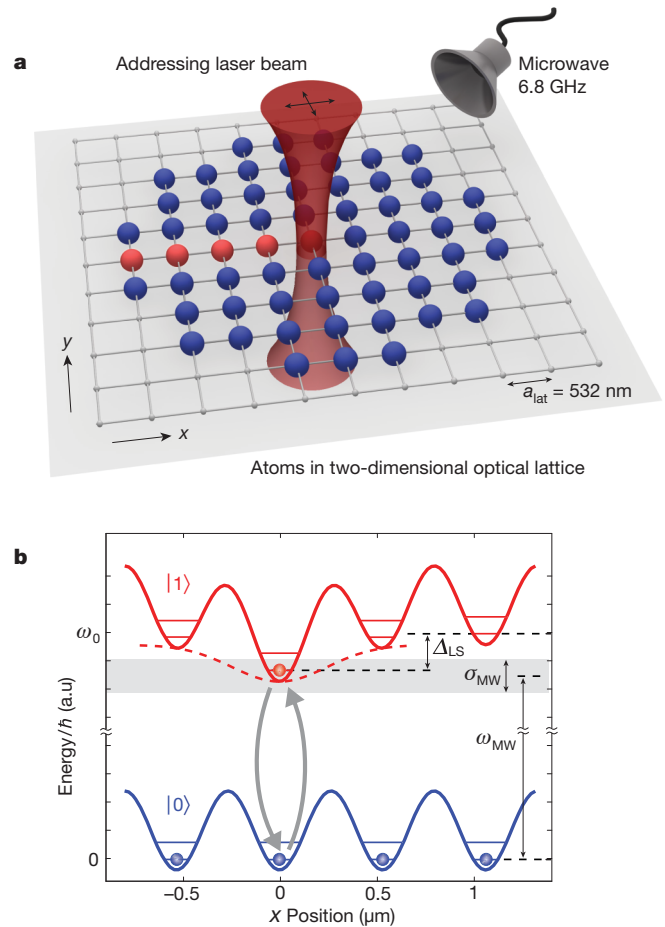


Figure 1 | Addressing scheme. **a**, Atoms in a Mott insulator with unity filling arranged on a square lattice with period $a_{\text{lat}} = 532$ nm were addressed using an off-resonant laser beam. The beam was focused onto individual lattice sites by a high-aperture microscope objective (not shown) and could be moved in the x - y plane with an accuracy of better than $0.1 a_{\text{lat}}$. **b**, Energy diagram of atoms in the lattice for the two hyperfine states $|0\rangle$ and $|1\rangle$ (energy separation $\hbar\omega_0$). The σ^- -polarized addressing beam locally induces a light shift Δ_{LS} of state $|1\rangle$, bringing it into resonance with a microwave field. A Landau-Zener sweep (central frequency ω_{MW} , sweep width σ_{MW}) transfers the addressed atoms from $|0\rangle$ to $|1\rangle$.

In order to reveal only the spin-flipped atoms, the spin states of all atoms were flipped by a global microwave sweep before the push-out laser was applied (Fig. 2c, d). However, owing to the finite transfer efficiency of the global sweep, some atoms remaining in state $|0\rangle$ were clearly visible in addition to the addressed ones. To avoid this problem when detecting the addressed atoms, we initially transferred the whole sample to state $|1\rangle$ by a microwave sweep and then shone in a repumping laser that completely depopulated state $|0\rangle$. Then, we used our addressing scheme to transfer selected atoms back to $|0\rangle$ and subsequently pushed out the atoms in $|1\rangle$, yielding typical images as shown in Fig. 2b, e, f.

Spin-flip fidelity

We quantified the success rate of our addressing scheme by again producing a series of spin-flips along the y lattice axis in a Mott insulator with unity filling (Fig. 3). The experimental procedure was the same as described above for the realization of Fig. 2a, in which the addressed sites were detected as empty sites. From the reconstructed atom number distribution (Fig. 3a), we determined the probability $p_0(\delta x)$ of finding an empty site as a function of the pointing offset δx between the addressing beam and the centre of the lattice site (Fig. 3b). We also investigated the effect of the addressing on neighbouring

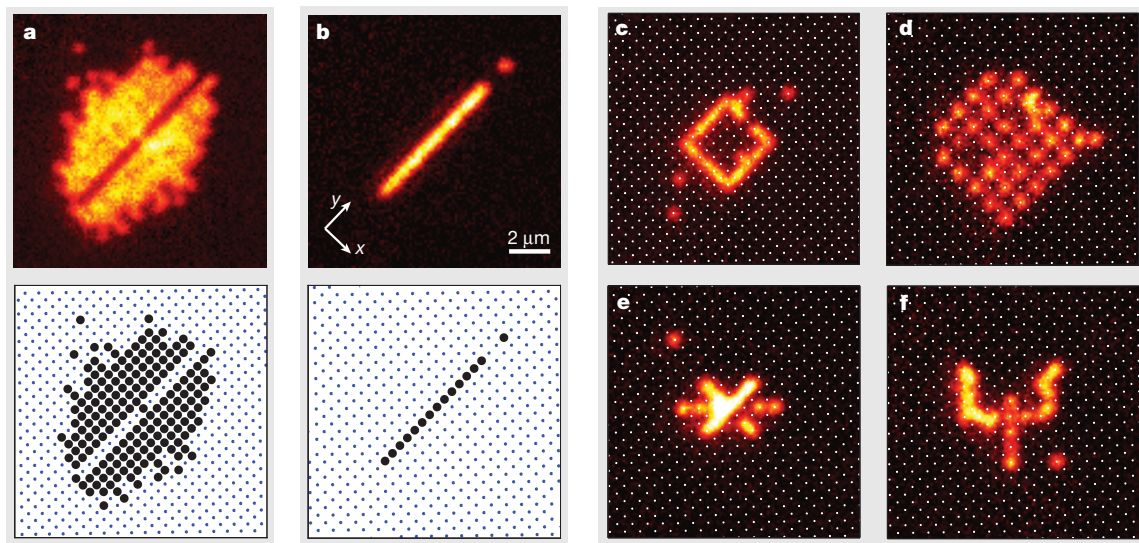


Figure 2 | Single-site addressing. **a**, Top, experimentally obtained fluorescence image of a Mott insulator with unity filling in which the spin of selected atoms was flipped from $|0\rangle$ to $|1\rangle$ using our single-site addressing scheme. Atoms in state $|1\rangle$ were removed by a resonant laser pulse before detection. Bottom, the reconstructed atom number distribution on the lattice. Each filled circle indicates a single atom; the points mark the lattice sites. **b**, Top, as for **a** except that a global microwave sweep exchanged the population in $|0\rangle$

and $|1\rangle$, such that only the addressed atoms were observed. Bottom, the reconstructed atom number distribution shows 14 atoms on neighbouring sites. **c–f**, As for **b**, but omitting the atom number distribution. The images contain 29 (**c**), 35 (**d**), 18 (**e**) and 23 (**f**) atoms. The single isolated atoms in **b**, **e** and **f** were placed intentionally to allow for the correct determination of the lattice phase for the feedback on the addressing beam position.

atoms, which ideally should remain unaffected. For this purpose, we monitored the probability of finding a hole at the sites next to the addressed ones (dark blue regions in Fig. 3a, b and points in Fig. 3c). In order to distinguish accidentally flipped neighbouring atoms from holes that originate from thermal excitations of the initial Mott insulator²⁸, we also monitored the probability of finding a hole at the second next neighbours (light blue regions and points in Fig. 3). As both yielded the same hole probability of 6(2)%, we attribute all holes to thermal excitations and conclude that the probability of addressing a neighbouring atom is indiscernibly small. We fitted the hole probability $p_0(\delta x)$ of the addressed site with a flat-top model function (see Methods), keeping the offset fixed at the thermal contribution of 6%. From the fit, we derived a spin-flip fidelity of 95(2)%, an FWHM of $\sigma_a = 330(10)$ nm and an edge sharpness of $\sigma_s = 50(10)$ nm (Fig. 3c). These values correspond to 60% and 10% of the addressing beam diameter, demonstrating that our method reaches sub-diffraction-limited resolution, well below the lattice spacing.

The observed maximum spin-flip fidelity is currently limited by the population transfer efficiency of our microwave sweep. The edge sharpness σ_s originates from the beam pointing error of $\lesssim 0.1 a_{\text{lat}}$ and from variations in the magnetic bias field. The latter causes frequency fluctuations of ~ 5 kHz, which translate into an effective pointing error of $0.05 a_{\text{lat}}$ at the maximum slope of the addressing beam profile. The resolution σ_a could in principle be further reduced by a narrower microwave sweep, at the cost of a larger sensitivity to the magnetic field fluctuations. A larger addressing beam power would reduce this sensitivity, but we observed that this deformed the lattice potential, owing to the imperfect σ^- -polarization, allowing neighbouring atoms to tunnel to the addressed sites.

Coherent tunnelling dynamics

The preparation of an arbitrary atom distribution opens up new possibilities for exploring coherent quantum dynamics at the single-atom level. As an example, we studied the tunnelling dynamics in a one-dimensional lattice (Fig. 4) which allowed us to determine how much our addressing scheme affects the vibrational state of the atoms. We started by preparing a single line of up to 18 atoms along the y direction before we lowered the lattice along the x direction to $V_x = 5.0(5) E_r$

within 200 μs . At the same time, the other lattices were lowered to $V_y = 30 E_r$ and $V_z = 23 E_r$, which reduced the external confinement along the x direction, but still suppressed tunnelling in the y and z directions. After a varying hold time t , allowing the atoms to tunnel along x , the atomic distribution was frozen by a rapid 100 μs ramp of all lattice axes to 56–90 E_r . By averaging the resulting atomic distribution along the y direction and repeating the experiment several times, we obtained the probability distribution of finding an atom at the different lattice sites (Fig. 4, bottom row).

This probability distribution samples the single-atom wavefunction after a coherent tunnelling evolution. We observed how the wavefunction expands in the lattice and how the interference of different paths leads to distinct maxima and minima in the distribution, leaving for example almost no atoms at the initial position after a single tunnelling time (Fig. 4c). This behaviour differs markedly from the evolution in free space, where a Gaussian wave packet disperses without changing its shape, always preserving a maximum probability in the centre. For longer hold times, an asymmetry in the spatial distribution becomes apparent (Fig. 4d), which originates from an offset between the bottom of the external harmonic confinement and the initial position of the atoms.

We describe the observed tunnelling dynamics by a simple Hamiltonian including the tunnel coupling $J^{(0)}$ between two neighbouring sites and an external harmonic confinement, parameterized by the trap frequency ω_{trap} , and the position offset x_{offs} (Methods). A single fit to all probability distributions recorded at different hold times yields $J^{(0)}/\hbar = 940(20)$ Hz, $\omega_{\text{trap}}/(2\pi) = 103(4)$ Hz and $x_{\text{offs}} = -6.3(6) a_{\text{lat}}$. This is in agreement with the trap frequency $\omega_{\text{trap}}/(2\pi) = 107(2)$ Hz obtained from an independent measurement via excitation of the dipole mode without the x lattice, whose contribution to the external confinement is negligible compared to the other two axes. From $J^{(0)}$, we calculated a lattice depth of $V_x = 4.6(1) E_r$, which agrees with an independent calibration via parametric heating. The expansion of the wave packet can also be understood by writing the initial localized wavefunction as a superposition of all Bloch waves of quasi-momentum $\hbar q$, with $-\pi/a_{\text{lat}} \leq q \leq \pi/a_{\text{lat}}$. To each quasi-momentum $\hbar q$, one can assign a velocity $v_q = \frac{1}{\hbar} \frac{\partial E}{\partial q}$, determined by the dispersion relation $E(q) = -2J^{(0)} \cos(qa_{\text{lat}})$ of the lowest band. The edges of the wave

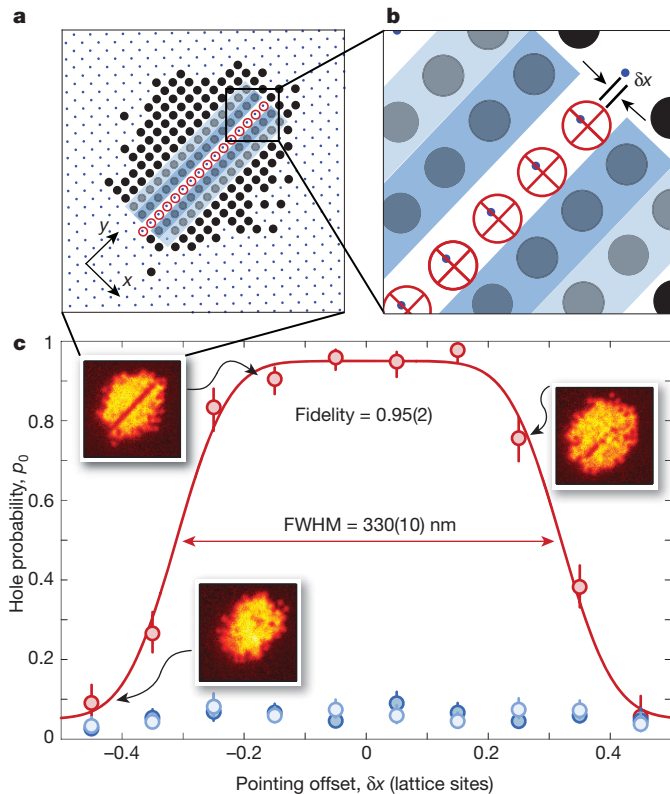


Figure 3 | Addressing fidelity. **a**, The spin-flip probability was measured by sequentially addressing a series of 16 neighbouring sites along the y lattice axis (red circles) in a Mott insulator with unity filling. **b**, Magnified view of boxed area in **a**, showing definition of pointing offset, δx . **c**, Plot of the resulting hole probability $p_0(\delta x)$ as a function of the pointing offset δx (red circles). Each point was obtained by averaging over 4–7 pictures (total 50–100 addressed lattice sites), taking only those sites into account which lie well within a Mott shell with unity filling. The displayed error bars show the 1σ statistical uncertainty, given by the Clopper-Pearson confidence limits. The data were fitted by a flat-top model function (Methods) and yields FWHM $\sigma_a = 330(20)$ nm, an edge sharpness of $\sigma_s = 50(10)$ nm, and a peak fidelity of 95(2)%. The offset was fixed at the 6(2)% probability of thermally activated holes, as deduced from the neighbouring and next neighbouring sites (dark and light blue shaded regions in **a** and **b**, and dark and light blue circles in **c**).

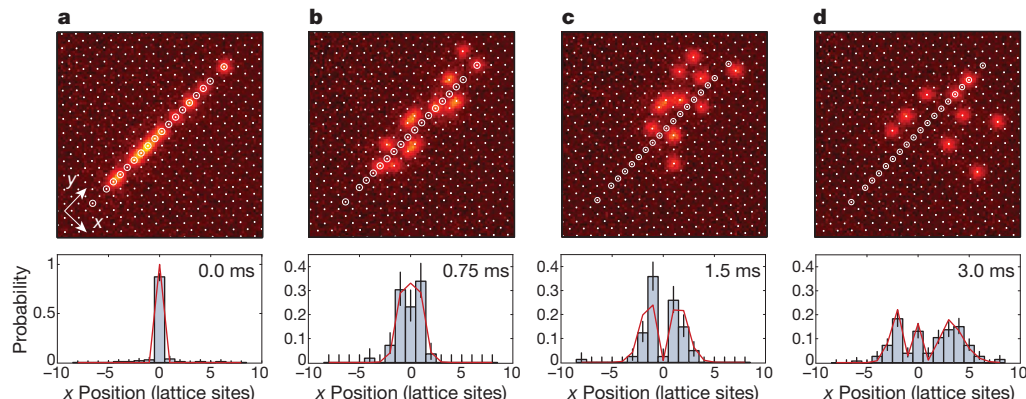


Figure 4 | Ground state tunnelling dynamics. **a**, Atoms were prepared in a single line along the y direction before the lattice along the x axis was lowered, allowing the atoms to tunnel in this direction (**b–d**). The top row shows fluorescence images of the atomic distribution after different hold times. White circles indicate the lattice sites at which the atoms were prepared (not all sites initially contained an atom). The bottom row shows the respective position

packet propagate with the largest occurring velocity $v_{\max} = 2f^{(0)}a_{\text{lat}}/\hbar = 1.88(4) a_{\text{lat}}/\text{ms}$, in agreement with our data.

Our measurements constitute the first observation of the ground state tunnelling dynamics of massive particles on a lattice with single-site resolution. A similar continuous-time quantum walk has been demonstrated with photons in an array of evanescently coupled photonic waveguides³⁷. For massive particles, a discrete quantum walk of single atoms has been observed using a sequence of spin manipulations and spin-dependent transports in an optical lattice³⁸ and also with trapped ions³⁹. Without single-particle and single-site resolution, a continuous-time quantum walk in the ground state has been observed for ultracold fermionic atoms by measuring their ballistic expansion in a weak lattice⁴⁰.

In a second tunnelling experiment, we observed the faster dynamics of atoms in the first excited band (Fig. 5). For this, we deliberately excited the atoms by introducing a pointing offset δx of the addressing beam, which caused a shift of the potential wells during the switch-on. We repeated the same tunnelling experiment as above with a hold time of $t = 1$ ms for different pointing offsets δx . For a small pointing offset ($\delta x = 0.1 a_{\text{lat}}$ in Fig. 5b) we observed a narrow distribution, compared to a much broader one for a large offset ($\delta x = -0.4 a_{\text{lat}}$ in Fig. 5a). We attribute this to different fractions f of atoms in the first band, which is characterized by the higher tunnelling rate $f^{(1)}$. We fitted the distribution of Fig. 5a to a two-band model (Methods) and found $f^{(1)}/\hbar = 6.22(6)$ kHz. This is in excellent agreement with the expected value of $f^{(1)}/\hbar = 6.14(6)$ kHz from a band structure calculation in which we used $f^{(0)}$ as an input parameter to calculate the lattice depth. Our measurement of the fraction of excited atoms f as a function of the pointing offset δx (Fig. 5b inset) shows that the atoms are strongly heated for large pointing offsets. By contrast, only a small fraction of the atoms is excited to the first band for small pointing offsets $|\delta x| \leq 0.1 a_{\text{lat}}$, yielding a ground state population of $1 - f = 87(2)\%$.

Discussion

In summary, we have demonstrated full two-dimensional single-site and single-atom spin control in an optical lattice with sub-diffraction-limited spatial resolution. Starting from a Mott insulator with unity filling, we achieved a spin-flip fidelity of 95(2)% with negligible influence on the neighbouring lattice sites. Our scheme leaves most of the atoms in the vibrational ground state. The control of single spins in a strongly correlated many-body system on a lattice opens many new possibilities for studying quantum dynamics and quantum phases. Our technique will allow us to create out-of-equilibrium states or local perturbations in order to observe the ensuing dynamics of the many-body

distribution obtained from an average over 10–20 of such pictures, the error bars give the 1σ statistical uncertainty. A single fit to all distributions recorded at different hold times (red curve) yields a tunnelling coupling of $f^{(0)}/\hbar = 940(20)$ Hz, a trap frequency of $\omega_{\text{trap}}/(2\pi) = 103(4)$ Hz and a trap offset of $x_{\text{offs}} = -6.3(6) a_{\text{lat}}$.

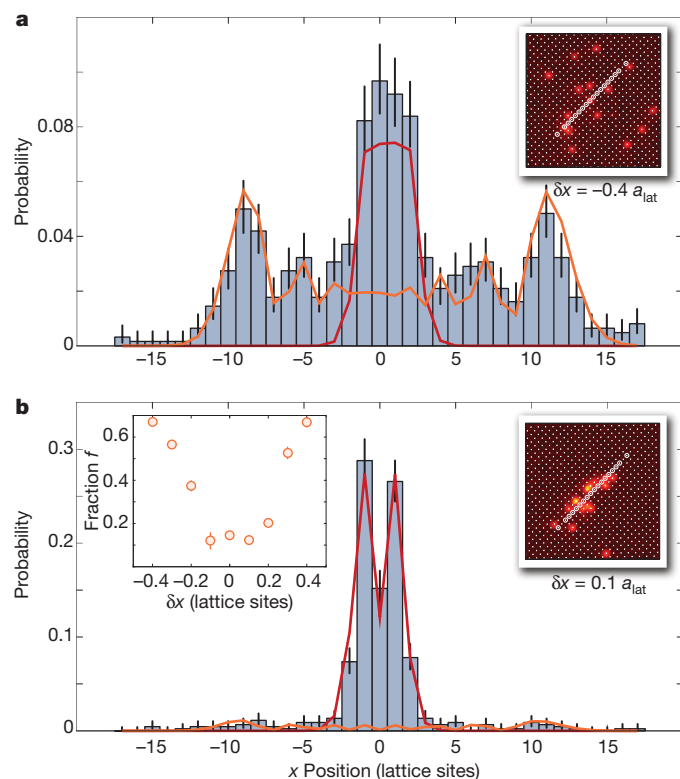


Figure 5 | Tunnelling dynamics of the first excited band. Some atoms were excited to the first band by a pointing offset δx of the addressing beam. **a, b**, Atomic position distribution after 1 ms tunnelling time for $\delta x = -0.4 a_{\text{lat}}$ (**a**) and $\delta x = 0.1 a_{\text{lat}}$ (**b**). Error bars give the 1σ statistical uncertainty. We fitted the data with a model that includes atoms in the zeroth band (red line) and a fraction f in the first band (orange line) (Methods). Right insets of **a, b** show corresponding original images. Left inset of **b** shows fraction f in first band versus pointing offset δx with a broad minimum of $f = 13(2)\%$, indicating that most of the atoms are left in the ground state.

system, such as spin-charge separation^{5,6} or spin impurity dynamics beyond Luttinger liquid theory⁴. Our studies of the tunnelling dynamics at the single-atom level can be extended to correlated particle tunnelling^{41–43}, also in higher dimensions, or to observe transport across local impurities⁴⁴ or potential barriers. The tunnelling can also be used to circumvent the pairwise losses during the imaging^{27,28} by letting the atoms of a one-dimensional system spread along the perpendicular direction in order to obtain a sufficiently low density. Further prospects are the implementation of novel cooling schemes relying on the local removal of regions with high entropy^{8,45}. The single-spin control in our large systems with several hundreds of atoms also opens new perspectives for scalable quantum computing. Combining single-qubit manipulation with local readout and a global entanglement operation in a spin-dependent lattice^{46,47} would be the basis of a one-way quantum computer^{11,12}. For the circuit model of a quantum computer, two-qubit operations can be realized by Rydberg gates between selected atom pairs in the lattice^{48,49}.

METHODS SUMMARY

The calibration of the addressing beam position was performed by replacing the far detuned addressing laser beam by a near resonant molasses beam. A fluorescence image of a large thermal atom cloud in the vertical lattice alone yielded a strongly enhanced signal at the position of the focused beam, which could be determined with an uncertainty of $0.05 a_{\text{lat}} \approx 25$ nm. In order to compensate slow phase drifts of the optical lattice, we applied a feedback on the position of the addressing beam. We determined the two lattice phases along x and y after each realization of the experiment by fitting the position of isolated atoms. Our microwave sweeps used for the spin-flips are HS1-pulses⁵⁰ of 20 ms duration, 60 kHz sweep width and -75 kHz offset of the sweep centre from the bare resonance. We

describe the coherent tunnelling dynamics by a Hamiltonian that includes nearest neighbour tunnel coupling in the lowest band, and an external harmonic confinement. The single particle wavefunction and its coherent time evolution are given by the Schrödinger equation. When including tunnelling in the first band, we assume an incoherent sum of the distributions of the zeroth and of the first band.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 December 2010; accepted 11 January 2011.

1. Binnig, G. & Rohrer, H. Scanning tunneling microscopy — from birth to adolescence. *Rev. Mod. Phys.* **59**, 615–625 (1987).
2. Giessibl, F. J. Advances in atomic force microscopy. *Rev. Mod. Phys.* **75**, 949–983 (2003).
3. Blatt, R. & Wineland, D. Entangled states of trapped atomic ions. *Nature* **453**, 1008–1015 (2008).
4. Zvonarev, M. B., Cheianov, V. V. & Giamarchi, T. Spin dynamics in a one-dimensional ferromagnetic Bose gas. *Phys. Rev. Lett.* **99**, 240404 (2007).
5. Recati, A., Fedichev, P. O., Zwerger, W. & Zoller, P. Spin-charge separation in ultracold quantum gases. *Phys. Rev. Lett.* **90**, 020401 (2003).
6. Kleine, A., Kollath, C., McCulloch, I. P., Giamarchi, T. & Schollwöck, U. Spin-charge separation in two-component Bose gases. *Phys. Rev. A* **77**, 013607 (2008).
7. Weiss, D. et al. Another way to approach zero entropy for a finite system of atoms. *Phys. Rev. A* **70**, 040302 (2004).
8. Bernier, J.-S. et al. Cooling fermionic atoms in optical lattices by shaping the confinement. *Phys. Rev. A* **79**, 061601(R) (2009).
9. Weimer, H., Müller, M., Lesanovsky, I., Zoller, P. & Büchler, H. P. A Rydberg quantum simulator. *Nature Phys.* **6**, 382–388 (2010).
10. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
11. Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Phys. Rev. Lett.* **86**, 5188–5191 (2001).
12. Briegel, H. J., Browne, D. E., Dür, W., Raussendorf, R. & Van Den Nest, M. Measurement-based quantum computation. *Nature Phys.* **5**, 19–26 (2009).
13. Dumke, R. et al. Micro-optical realization of arrays of selectively addressable dipole traps: a scalable configuration for quantum computation with atomic qubits. *Phys. Rev. Lett.* **89**, 097903 (2002).
14. Bergamini, S. et al. Holographic generation of microtrap arrays for single atoms by use of a programmable phase modulator. *J. Opt. Soc. Am. B* **21**, 1889–1894 (2004).
15. Saffman, M. Addressing atoms in optical lattices with Bessel beams. *Opt. Lett.* **29**, 1016–1018 (2004).
16. Calarco, T., Dörner, U., Julienne, P. S., Williams, C. J. & Zoller, P. Quantum computations with atoms in optical lattices: marker qubits and molecular interactions. *Phys. Rev. A* **70**, 012306 (2004).
17. Zhang, C., Rolston, S. L. & Das Sarma, S. Manipulation of single neutral atoms in optical lattices. *Phys. Rev. A* **74**, 042316 (2006).
18. Joo, J., Lim, Y. L., Beige, A. & Knight, P. L. Single-qubit rotations in two-dimensional optical lattices with multiqubit addressing. *Phys. Rev. A* **74**, 042344 (2006).
19. Cho, J. Addressing individual atoms in optical lattices with standing-wave driving fields. *Phys. Rev. Lett.* **99**, 020502 (2007).
20. Gorshkov, A. V., Jiang, L., Greiner, M., Zoller, P. & Lukin, M. D. Coherent quantum optical control with subwavelength resolution. *Phys. Rev. Lett.* **100**, 093005 (2008).
21. Lundblad, N., Obrecht, J. M., Spielman, I. B. & Porto, J. V. Field-sensitive addressing and control of field-insensitive neutral-atom qubits. *Nature Phys.* **5**, 575–580 (2009).
22. Shibata, K., Kato, S., Yamaguchi, A., Uetake, S. & Takahashi, Y. A scalable quantum computer with ultranarrow optical transition of ultracold neutral atoms in an optical lattice. *Appl. Phys. B* **97**, 753–758 (2009).
23. Scheunemann, R., Cataliotti, F. S., Hänsch, T. W. & Weitz, M. Resolving and addressing atoms in individual sites of a CO₂-laser optical lattice. *Phys. Rev. A* **62**, 051801(R) (2000).
24. Schrader, D. et al. Neutral atom quantum register. *Phys. Rev. Lett.* **93**, 150501 (2004).
25. Karski, M. et al. Imprinting patterns of neutral atoms in an optical lattice using magnetic resonance techniques. *N. J. Phys.* **12**, 065027 (2010).
26. Würtz, P., Langen, T., Gericke, T., Koglbauer, A. & Ott, H. Experimental demonstration of single-site addressability in a two-dimensional optical lattice. *Phys. Rev. Lett.* **103**, 080404 (2009).
27. Bakr, W. S. et al. Probing the superfluid-to-Mott insulator transition at the single-atom level. *Science* **329**, 547–550 (2010).
28. Sherson, J. F. et al. Single-atom-resolved fluorescence imaging of an atomic Mott insulator. *Nature* **467**, 68–72 (2010).
29. Schlosser, N., Reymond, G., Protchenko, I. & Grangier, P. Sub-poissonian loading of single atoms in a microscopic dipole trap. *Nature* **411**, 1024–1027 (2001).
30. Kuhr, S. et al. Deterministic delivery of a single atom. *Science* **293**, 278–280 (2001).
31. Grünzweig, T., Hilliard, A., McGovern, M. & Andersen, M. F. Near-deterministic preparation of a single atom in an optical microtrap. *Nature Phys.* **6**, 951–954 (2010).
32. Fisher, M. P. A., Weichman, P. B., Grinstein, G. & Fisher, D. S. Boson localization and the superfluid-insulator transition. *Phys. Rev. B* **40**, 546–570 (1989).

33. Jaksch, D., Bruder, C., Cirac, J. I., Gardiner, C. & Zoller, P. Cold bosonic atoms in optical lattices. *Phys. Rev. Lett.* **81**, 3108–3111 (1998).
34. Greiner, M., Mandel, O., Esslinger, T., Hänsch, T. W. & Bloch, I. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature* **415**, 39–44 (2002).
35. Fölling, S., Widera, A., Müller, T., Gerbier, F. & Bloch, I. Formation of spatial shell structure in the superfluid to Mott insulator transition. *Phys. Rev. Lett.* **97**, 060403 (2006).
36. Campbell, G. K. *et al.* Imaging the Mott insulator shells by using atomic clock shifts. *Science* **313**, 649–652 (2006).
37. Perets, H. B. *et al.* Realization of quantum walks with negligible decoherence in waveguide lattices. *Phys. Rev. Lett.* **100**, 170506 (2008).
38. Karski, M. *et al.* Quantum walk in position space with single optically trapped atoms. *Science* **325**, 174–177 (2009).
39. Zähringer, F. *et al.* Realization of a quantum walk with one and two trapped ions. *Phys. Rev. Lett.* **104**, 100503 (2010).
40. Schneider, U. *et al.* Breakdown of diffusion: from collisional hydrodynamics to a continuous quantum walk in a homogeneous Hubbard model. Preprint at (<http://arXiv.org/abs/1005.3545v1>) (2010).
41. Winkler, K. *et al.* Repulsively bound atom pairs in an optical lattice. *Nature* **441**, 853–856 (2006).
42. Fölling, S. *et al.* Direct observation of second-order atom tunnelling. *Nature* **448**, 1029–1032 (2007).
43. Peruzzo, A. *et al.* Quantum walks of correlated photons. *Science* **329**, 1500–1503 (2010).
44. Micheli, A., Daley, A. J., Jaksch, D. & Zoller, P. Single atom transistor in a 1D optical lattice. *Phys. Rev. Lett.* **93**, 140408 (2004).
45. Sherson, J. F. & Mølmer, K. Shaking the entropy out of a lattice: atomic filtering by vibrational excitations. Preprint at (<http://arXiv.org/abs/1012.1457v1>) (2010).
46. Jaksch, D., Briegel, H.-J., Cirac, J. I., Gardiner, C. W. & Zoller, P. Entanglement of atoms via cold controlled collisions. *Phys. Rev. Lett.* **82**, 1975–1978 (1999).
47. Mandel, O. *et al.* Controlled collisions for multiparticle entanglement of optically trapped atoms. *Nature* **425**, 937–940 (2003).
48. Wilk, T. *et al.* Entanglement of two individual neutral atoms using Rydberg blockade. *Phys. Rev. Lett.* **104**, 010502 (2010).
49. Isenhower, L. *et al.* Demonstration of a neutral atom controlled-NOT quantum gate. *Phys. Rev. Lett.* **104**, 010503 (2010).
50. Garwood, M. & Delabarre, L. The return of the frequency sweep: designing adiabatic pulses for contemporary NMR. *J. Magn. Reson.* **153**, 155–177 (2001).

Acknowledgements We thank W. Ketterle for discussions and ideas. We acknowledge the help of R. Glöckner and R. Labouvie during the construction of the experiment. We acknowledge funding by MPG, DFG, Stiftung Rheinland-Pfalz für Innovation, Carl-Zeiss Stiftung, EU (NAMEQUAM, AQUATE, Marie Curie Fellowships to J.F.S. and M.C.), and JSPS (Postdoctoral Fellowship for Research Abroad to T.F.).

Author Contributions All authors contributed to the acquisition and analysis of the data; C.W., M.E., J.F.S., M.C. and S.K. designed and constructed the apparatus; C.W., I.B. and S.K. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.K. (stefan.kuhr@mpq.mpg.de).

METHODS

Calibration of the addressing beam position. To move the addressing laser beam in the object plane, we changed the angle of the beam entering from the reverse direction into the microscope objective using a two-axis piezo mirror. The device has an angular resolution of $5 \mu\text{rad}$, corresponding to a theoretical resolution in the object plane of $0.02 a_{\text{lat}} \approx 10 \text{ nm}$. In order to position the addressing laser beam onto the atoms with high precision, we measured calibration functions that translate the two control voltages of the piezo mirror into image coordinates. This calibration was performed by replacing the far detuned addressing laser beam by a near resonant molasses beam that follows the identical beam path. Using in addition the x and y molasses beams, we took a fluorescence image of a large thermal atom cloud in the vertical lattice alone and observed a strongly enhanced signal at the position of the focused beam. We determined the position of this fluorescence maximum with an uncertainty of 0.2 pixels in our images, corresponding to $0.05 a_{\text{lat}} \approx 25 \text{ nm}$ in the object plane. The long term drifts of the addressing beam position are on the order of $0.1 a_{\text{lat}}$ per hour, which we took into account by regular recalibration of the beam position.

Lattice phase feedback. In order to compensate slow phase drifts of the optical lattice, we applied a feedback on the position of the addressing beam. We determined the two lattice phases along x and y after each realization of the experiment by fitting the position of isolated atoms. Averaging over the positions of typically 1–5 isolated atoms per image allowed us to determine the lattice phase to better than $0.05 a_{\text{lat}}$. For the determination of the phase, we used the lattice constant and the lattice angles determined from a fluorescence image with many isolated atoms²⁸. As our phase drifts were slower than $0.04 a_{\text{lat}}$ between two successive realizations of the experiment (25 s cycle time), we used the lattice phase from the previous image to correct the addressing beam position. This feedback was done by adding appropriate offsets to the piezo control voltages.

Microwave sweeps. Our microwave sweeps are HS1-pulses⁵⁰ with time-dependent Rabi frequency $\Omega(t)$ and detuning $\delta(t)$ given by

$$\Omega(t) = \Omega_0 \operatorname{sech} \left[\beta \left(\frac{2t}{T_p} - 1 \right) \right] \quad (1)$$

$$\delta(t) = \frac{\sigma_{\text{MW}}}{2} \tanh \left[\beta \left(\frac{2t}{T_p} - 1 \right) \right] \quad (2)$$

where $\Omega_0/(2\pi) = 3 \text{ kHz}$ is the maximum Rabi frequency, $\beta = 5$ is a truncation factor, $T_p = 20 \text{ ms}$ is the pulse length, and $\sigma_{\text{MW}}/(2\pi) = 60 \text{ kHz}$ is the sweep width. The detuning $\delta(t)$ is measured relative to the centre of the sweep at $\omega_{\text{MW}} = \omega_0 - \Delta_{\text{MW}}$ (see Fig. 1b). Here, ω_0 is the bare resonance between the two hyperfine states, including the shift of -570 kHz due to the magnetic bias field along the z direction, and $\Delta_{\text{MW}}/(2\pi) = -75 \text{ kHz}$ denotes the offset of the sweep centre.

Spin-flip fidelity. In order to determine the spin-flip fidelity, we fitted the hole probability p_0 as a function of the pointing offset δx (Fig. 3b) to a flat-top model function given by:

$$p_0(\delta x) = \frac{A}{2} \left[\operatorname{erf} \left(\frac{\delta x + \sigma_a/2}{\sigma_s} \right) + \operatorname{erf} \left(-\frac{\delta x - \sigma_a/2}{\sigma_s} \right) \right] + B \quad (3)$$

Here, $\operatorname{erf}(x) = 2/\sqrt{\pi} \int_0^x e^{-\tau^2} d\tau$ is the error function, σ_a denotes the full-width at half-maximum of the flat-top profile and σ_s the edge sharpness. We chose this model function since our HS1-pulses (see above) produce a flat-top population transfer profile, the edges of which are dominated by randomly fluctuating quantities (beam pointing and magnetic fields) following Gaussian statistics. The addressing fidelity is defined as $F = A/(1 - B)$ taking into account that the maximum hole probability $p_0^{\text{max}} = A + B$ also includes holes from thermal defects. These yield a hole with probability B at unsuccessfully addressed sites which occur with probability $1 - F$, such that $p_0^{\text{max}} = F + (1 - F)B$.

Single-particle tunnelling dynamics. We describe the coherent tunnelling dynamics on $k = 2n + 1$ lattice sites by the Hamiltonian

$$\hat{H}^{(0)} = -J^{(0)} \sum_{i=-n}^n (\hat{a}_{i+1}^\dagger \hat{a}_i + \hat{a}_{i-1}^\dagger \hat{a}_i) + V_{\text{ext}} \sum_{i=-n}^n (i - x_{\text{offs}})^2 \hat{a}_i^\dagger \hat{a}_i \quad (4)$$

where $J^{(0)}$ is the tunnel coupling in the lowest band, and \hat{a}_i^\dagger (\hat{a}_i) is the creation (annihilation) operator for a particle at site i . The strength of the external harmonic potential with trapping frequency ω_{trap} is given by $V_{\text{ext}} = \frac{1}{2} m \omega_{\text{trap}}^2 a_{\text{lat}}^2$, and x_{offs} describes a position offset with respect to the bottom of the harmonic potential. The single particle wavefunction and its coherent time evolution are given by

$$\Psi^{(0)}(t) = \sum_{i=-n}^n c_i^{(0)}(t) \hat{a}_i^\dagger |\tilde{0}\rangle = \exp(-i\hat{H}^{(0)}t/\hbar) \Psi^{(0)}(0) \quad (5)$$

with the initial condition $\Psi^{(0)}(0) = a_0^+ |\tilde{0}\rangle$ and the vacuum state $|\tilde{0}\rangle$. The resulting probability of finding the particle at lattice site i after time t is $P_i^{(0)}(t) = |c_i^{(0)}(t)|^2$. For analysing the data of Fig. 4, we calculated the time evolution for $k = 17$ lattice sites.

Tunnelling in the first excited band. When including tunnelling in the first band, we assume an incoherent sum $P^{\text{tot}}(t)$ of the distributions $P^{(0)}(t)$ of the zeroth and $P^{(1)}(t)$ of the first band as:

$$P^{\text{tot}}(t) = (1 - f)P^{(0)}(t) + fP^{(1)}(t) \quad (6)$$

The Hamiltonian $\hat{H}^{(1)}$ in the first band and the coherent dynamics are identical to the ones of the zeroth band (equations (4) and (5)), except for a different tunnel coupling $J^{(1)}$. When fitting this model to our data, we kept ω_{trap} , x_{offs} and $J^{(0)}$ fixed at the values obtained from the data displayed in Fig. 4. We extracted $J^{(1)}$ from the data of Fig. 5a and used this value to fit the results for other pointing offsets. For the data in Fig. 5, the parameters of our microwave sweep were such that neighbouring atoms were also addressed. We took this into account by summing over two distinct probability distributions with a second starting position in the direction of the pointing offset.

DICER1 deficit induces *Alu* RNA toxicity in age-related macular degeneration

Hiroki Kaneko^{1*}, Sami Dridi^{1*}, Valeria Tarallo^{1*}, Bradley D. Gelfand¹, Benjamin J. Fowler¹, Won Gil Cho^{1,2}, Mark E. Kleinman¹, Steven L. Ponicsan³, William W. Hauswirth⁴, Vince A. Chiodo⁴, Katalin Karikó⁵, Jae Wook Yoo⁶, Dong-ki Lee⁶, Majda Hadziahmetovic⁷, Ying Song⁷, Smita Misra⁸, Gautam Chaudhuri⁸, Frank W. Buaas⁹, Robert E. Braun⁹, David R. Hinton¹⁰, Qing Zhang¹¹, Hans E. Grossniklaus¹¹, Jan M. Provis¹², Michele C. Madigan^{13,14}, Ann H. Milam⁷, Nikki L. Justice¹, Romulo J. C. Albuquerque¹, Alexander D. Blandford¹, Sasha Bogdanovich¹, Yoshio Hirano¹, Jassir Witta¹⁵, Elaine Fuchs¹⁶, Dan R. Littman¹⁷, Balamurali K. Ambati^{18,19}, Charles M. Rudin²⁰, Mark M. W. Chong^{17,21}, Patrick Provost²², Jennifer F. Kugel³, James A. Goodrich³, Joshua L. Dunaief⁷, Judit Z. Baffi¹ & Jayakrishna Ambati^{1,23}

Geographic atrophy (GA), an untreatable advanced form of age-related macular degeneration, results from retinal pigmented epithelium (RPE) cell degeneration. Here we show that the microRNA (miRNA)-processing enzyme DICER1 is reduced in the RPE of humans with GA, and that conditional ablation of *Dicer1*, but not seven other miRNA-processing enzymes, induces RPE degeneration in mice. *DICER1* knockdown induces accumulation of *Alu* RNA in human RPE cells and *Alu*-like B1 and B2 RNAs in mouse RPE. *Alu* RNA is increased in the RPE of humans with GA, and this pathogenic RNA induces human RPE cytotoxicity and RPE degeneration in mice. Antisense oligonucleotides targeting *Alu*/B1/B2 RNAs prevent *DICER1* depletion-induced RPE degeneration despite global miRNA downregulation. *DICER1* degrades *Alu* RNA, and this digested *Alu* RNA cannot induce RPE degeneration in mice. These findings reveal a miRNA-independent cell survival function for *DICER1* involving retrotransposon transcript degradation, show that *Alu* RNA can directly cause human pathology, and identify new targets for a major cause of blindness.

Age-related macular degeneration (AMD), which is as prevalent as cancer in industrialized countries, is a leading cause of blindness^{1,2}. In contrast to neovascular AMD, the more common atrophic form of AMD is without effective therapy^{3,4}. Extensive atrophy of the retinal pigment epithelium (RPE) leads to severe vision loss and is termed GA, whose pathogenesis is unclear. Here, we identify dysregulation of the RNase *DICER1* (ref. 5) and the resulting accumulation of transcripts of *Alu* elements, the most common small interspersed repetitive elements in the human genome⁶, as a potential cause of GA, and demonstrate strategies to inhibit this pathology *in vivo*.

DICER1 loss in GA induces RPE death

In human donor eyes with GA ($n = 10$), *DICER1* mRNA abundance was reduced in the macular RPE by $65 \pm 3\%$ (mean \pm s.e.m.; $P = 0.0036$; Mann-Whitney *U*-test) compared to control eyes ($n = 11$) (Fig. 1a). In contrast, there was no change in the abundance of *DROSHA* and *DGCR8* mRNAs, whose gene products form a complex that processes pri-miRNAs into pre-miRNAs⁷, or of the gene encoding Argonaute 2 (AGO2, encoded by *EIF2C2*), the core

component of the miRNA effector complex^{8,9}. *DICER1* protein expression was reduced in the RPE, but not the neural retina, of eyes with GA compared to controls (Fig. 1b, c and Supplementary Figs 1 and 2).

Because *DICER1* is downregulated in chemically stressed cells⁶, we tested whether *DICER1* reduction is common to dying retina. *DICER1* protein levels were not reduced in the RPE of human eyes with other retinal diseases (vitelliform macular dystrophy, retinitis pigmentosa, retinal detachment; Supplementary Fig. 3). Also, *Dicer1* abundance in the RPE was not reduced in numerous mouse models of retinal degeneration including *Ccl2*^{-/-} *Ccr2*^{-/-} (refs 7,8) and *Cp*^{-/-} *Heph*^{-/-} mice⁹ (Supplementary Fig. 3; Supplementary Notes). These data indicate that *DICER1* depletion in the RPE of eyes with GA is not a generic damage response.

To determine the consequence of *DICER1* reduction in the RPE, we interbred *Dicer1*^{fl/fl} mice¹⁰ with *BEST1* Cre mice¹¹, which express Cre recombinase under the control of the RPE cell-specific *BEST1* promoter. *BEST1* Cre; *Dicer1*^{fl/fl} mice uniformly showed RPE cell degeneration whereas littermate controls did not (Fig. 1d–f). We also

¹Department of Ophthalmology & Visual Sciences, University of Kentucky, Lexington, Kentucky 40506, USA. ²Department of Anatomy, Yonsei University Wonju College of Medicine, Wonju City 220-701, Korea. ³Department of Chemistry and Biochemistry, University of Colorado at Boulder, Boulder, Colorado 80309, USA. ⁴Department of Ophthalmology, University of Florida, Gainesville, Florida 32610, USA. ⁵Department of Neurosurgery, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. ⁶Global Research Laboratory for RNAi Medicine & BK21 School of Chemical Materials Science and Department of Chemistry, Sungkyunkwan University, Suwon 440-746, Korea. ⁷F.M. Kirby Center for Molecular Ophthalmology, Scheie Eye Institute, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. ⁸Department of Microbiology and Immunology, Meharry Medical College, Nashville, Tennessee 37208, USA. ⁹The Jackson Laboratory, Bar Harbor, Maine 04609, USA. ¹⁰The Arnold and Mabel Beckman Macular Research Center at the Doheny Eye Institute, University of Southern California, Los Angeles, California 90033, USA. ¹¹Departments of Ophthalmology & Pathology, Emory University Atlanta, Georgia 30322, USA. ¹²ARC Centre of Excellence in Vision Science, ANU Medical School and Research School of Biology, The Australian National University, Canberra, Australian Capital Territory 0200, Australia. ¹³School of Optometry and Vision Science, The University of New South Wales, Kensington, New South Wales 2033, Australia. ¹⁴Save Sight Institute, The University of Sydney, Sydney, New South Wales 2001, Australia. ¹⁵Department of Internal Medicine, University of Kentucky, Lexington, Kentucky 40506, USA. ¹⁶Howard Hughes Medical Institute, Laboratory of Mammalian Cell Biology and Development, The Rockefeller University, New York, New York 10065, USA. ¹⁷Howard Hughes Medical Institute, The Kimmel Center for Biology and Medicine of the Skirball Institute, New York University School of Medicine, New York, New York 10016, USA. ¹⁸Department of Ophthalmology and Visual Sciences, Moran Eye Center, University of Utah School of Medicine, Salt Lake City, Utah 84132, USA. ¹⁹Department of Ophthalmology, Veterans Affairs Salt Lake City Healthcare System, Salt Lake City, Utah 84148, USA. ²⁰Department of Oncology, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Johns Hopkins University, Baltimore, Maryland, 21231, USA. ²¹The Walter and Eliza Hall Institute, Autoimmunity and Transplantation Division, Parkville, Victoria 3052, Australia. ²²CHUL Research Center/CHUQ and Faculty of Medicine, Université Laval, Quebec, Quebec G1K 7P4, Canada. ²³Department of Physiology, University of Kentucky, Lexington, Kentucky 40506, USA.

*These authors contributed equally to this work.

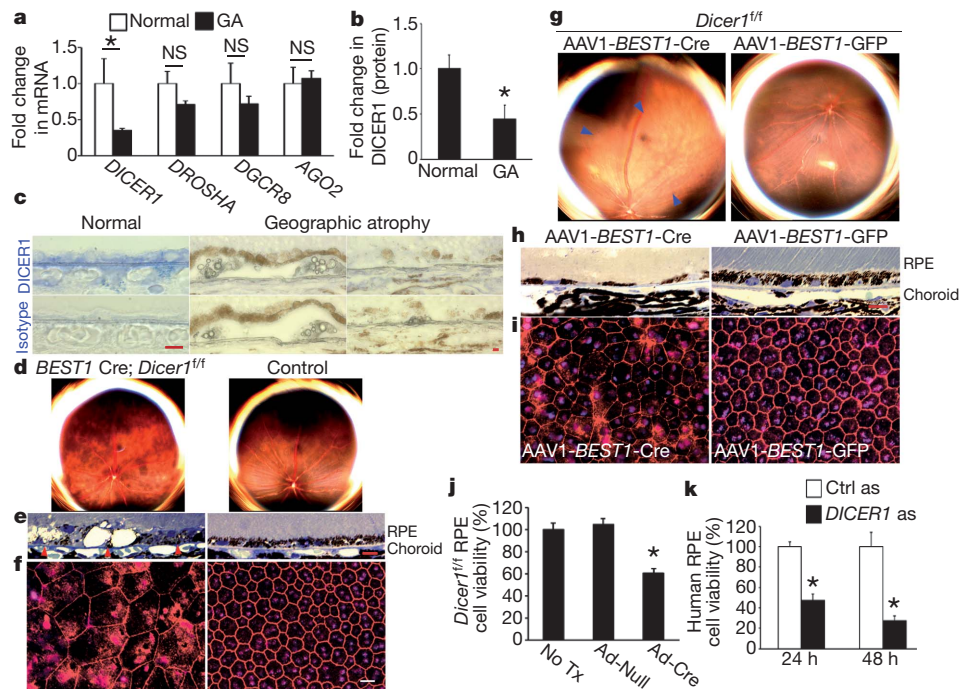


Figure 1 | *DICER1* deficit in GA induces RPE degeneration. **a**, *DICER1* is less abundant in RPE of human eyes with GA ($n = 10$) compared to control RPE ($n = 11$). $P = 0.004$ by Mann–Whitney U -test. *DROSHA*, *DGCR8* and *EIF2C2* (encoding *AGO2*) abundance not significantly different ($P > 0.11$ by Mann–Whitney U -test). $n = 10$ – 11 . **b**, *DICER1* quantification, assessed by western blotting (Supplementary Fig. 1), lower in human GA RPE ($n = 4$) compared to control RPE ($n = 4$). $P = 0.003$ by Student t -test. **c**, Reduced *DICER1* (blue) in human GA RPE compared to control eyes. **d**, **e**, Fundus photographs (**d**) and toluidine-blue-stained sections (**e**) show RPE degeneration in *BEST1* Cre; *Dicer1*^{+/f} mice but not controls. Arrowheads point to basal surface of RPE. **f**, Flat mounts stained for zonula occludens-1 (ZO-1; red) show RPE disruption in *BEST1* Cre; *Dicer1*^{+/f} mice compared to controls.

deleted *Dicer1* in adult mouse RPE by subretinal injection of an adeno-associated viral vector (AAV) coding for Cre recombinase under the control of the *BEST1* promoter¹² (AAV1-BEST1-Cre) in *Dicer1*^{+/f} mice (Supplementary Fig. 4). These eyes uniformly developed RPE cell degeneration, whereas contralateral eyes that underwent subretinal injection of AAV1-BEST1-GFP (GFP, green fluorescent protein) and wild-type mouse eyes injected with subretinal AAV1-BEST1-Cre did not (Fig. 1g–i and Supplementary Fig. 4). RPE cell dysmorphology in *Dicer1*-depleted mice resembled that of human GA eyes (Supplementary Fig. 5). When *Dicer1*^{+/f} mouse RPE cells were infected with an adenoviral vector coding for Cre recombinase (Ad-Cre), cell viability was reduced (Fig. 1j). Similarly, antisense oligonucleotide mediated knockdown of *DICER1* in human RPE cells increased cell death (Fig. 1k). Collectively, these data indicate that *DICER1* dysregulation is involved in the pathogenesis of GA.

DICER1 phenotype not due to miRNA dysregulation

We tested whether depleting other miRNA-processing enzymes induces RPE degeneration. Subretinal injection of AAV1-BEST1-Cre in *Drosha*^{+/f} (ref. 13), *Dgcr8*^{+/f} (refs. 13,14), or *Ago2*^{+/f} mice¹⁵ did not damage the RPE (Supplementary Fig. 6), indicating that miRNA imbalances are not responsible for RPE degeneration induced by *DICER1* depletion. However, some miRNAs are generated by *Dicer1* independent of *Drosha* and *Dgcr8* (refs 16, 17). There is also debate whether *Ago2* is essential for miRNA function^{15,18–21}. Mice deficient in *Ago1*, *Ago3*, or *Ago4* (also known as *Eif2c1*, *Eif2c3* and *Eif2c4*, respectively) had normal RPE (Supplementary Fig. 7). TRBP (encoded by *Tarbp2*) recruits *DICER1* to the four Argonaute proteins to enable miRNA processing and RNA silencing (ref. 22 and R.

Shiekhattar, personal communication); *Tarbp2*^{−/−} mice also had no RPE degeneration (Supplementary Fig. 7). These data indicate that RPE degeneration induced by *Dicer1* ablation involves a mechanism specific to *Dicer1* and not to miRNA machinery in general. To investigate further whether miRNA imbalances might contribute to the *DICER1* depletion phenotype, we studied human HCT116 colon cancer cells in which the helicase domain in exon 5 of *DICER1* was disrupted. Despite impaired miRNA biogenesis in these HCT-DICER1^{ex5} cells²³, baseline cell viability was not different between HCT-DICER1^{ex5} and parent HCT116 cells (Supplementary Fig. 8). These findings indicate that the principal biological effect of *DICER1* deficit contributing to the development of GA is not miRNA dysregulation, but do not exclude miRNA dysregulation promoting GA through other pathways.

Alu RNA accumulation in GA

Because miRNA perturbations were not implicated, we speculated that impaired processing of other dsRNAs might be involved. Using an antibody^{24,25} that recognizes long dsRNA (J2), we detected abundant dsRNA immunoreactivity in the macular RPE of human eyes with GA ($n = 10$; Fig. 2a–c) but not in control eyes ($n = 10$; Fig. 2d). We immunoprecipitated RPE lysates with J2 antibody and then sequenced the dsRNA using a T4 RNA ligase-aided, adaptor-based PCR amplification strategy. Approximately 300-nucleotide-long dsRNA species were found in the macular RPE of human eyes with GA (12/12) but not in eyes without GA (0/18) ($P = 1.2 \times 10^{-8}$ by Fisher's exact test) (Fig. 2e).

Alu RNA accumulation in GA

We recovered clones from 8 of the 12 GA eyes and identified two distinct sequences with high similarity ($E = 3.3 \times 10^{-103}$; 1.1×10^{-76})

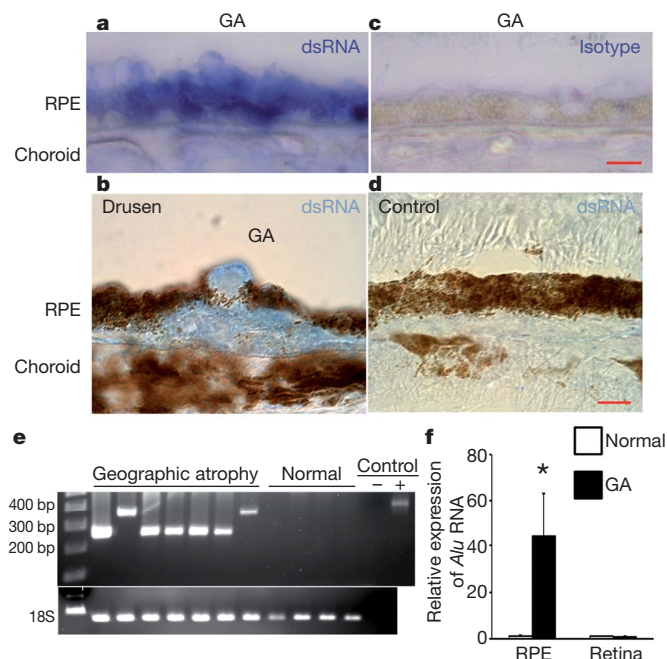


Figure 2 | *Alu* RNA accumulation in GA triggered by *DICER1* reduction. **a**, **b**, dsRNA immunolocalized (blue) in RPE (**a**, **b**) and sub-RPE deposits (drusen; **b**) in human GA. **c**, **d**, No staining with isotype antibody in GA RPE (**c**) and with anti-dsRNA antibody in control eye (**d**). Scale bars (**a**–**d**), 10 μ m. $n = 10$ (**a**–**d**). **e**, PCR amplification of immunoprecipitated dsRNA yielded amplicons with homology to *Alu* in GA RPE but not normal RPE. Water control (–) showed no amplification and recombinant dsRNA (+) showed predicted amplicon. **f**, Increased *Alu* RNA in GA RPE compared to control ($n = 7$). $P < 0.05$ by Student *t*-test. No significant difference in *Alu* RNA in neural retina. Values normalized to abundance in normal eyes.

to *Alu* RNAs (Supplementary Fig. 9). These sequences showed homology to the *Alu* Sq subfamily consensus sequence. *Alu* RNAs were the only dsRNA transcripts identified specifically in the J2-immunoprecipitated GA samples. We confirmed that J2 recognized *Alu* RNA (Supplementary Fig. 10). There was a marked increase in the abundance of *Alu* RNAs in the RPE of human eyes with GA compared to control eyes ($n = 7$), but not in the neural retina (Fig. 2f and Supplementary Fig. 11). The reference genome did not contain exact matches to these *Alu* sequences. This could be attributed to genetic variations or regions not represented in the reference genome or to chimaeric *Alu* formation. Further studies should elucidate the genomic origin of and regulatory factors involved in transcription of these *Alu* RNAs.

DICER1 depletion induces *Alu* RNA cytotoxicity

We tested whether *Alu* RNA accumulation in the RPE of GA was due to deficient *DICER1* processing. *DICER1* knockdown in human RPE cells using antisense oligonucleotides increased *Alu* RNA accumulation (Fig. 3a and Supplementary Fig. 12). Ad-Cre infection of *Dicer1*^{fl/fl} mouse RPE cells induced accumulation of B1 and B2 RNAs (Fig. 3b, c). *DICER1* was expressed in the nucleus and cytoplasm of RPE cells and its depletion induced accumulation of *Alu*/B1/B2 RNA in both compartments (Fig. 3b–d and Supplementary Fig. 13). Recombinant *DICER1*, but not heat-denatured *DICER1*, degraded *Alu* RNA (Fig. 3e). Enforced expression of *DICER1* in human RPE cells reduced the abundance of overexpressed *Alu* RNA (Fig. 3f), consistent with their degradation by *DICER1* *in vivo*. These data confirm that *DICER1* dysregulation can trigger *Alu*/B1/B2 RNA accumulation.

Because cell stresses can induce generalized retrotransposon activation, we wondered whether *Alu* RNA accumulation in GA might be a generic response in dying retina. However, in the RPE of human eyes with GA and in *DICER1*-depleted human RPE cells, there was no

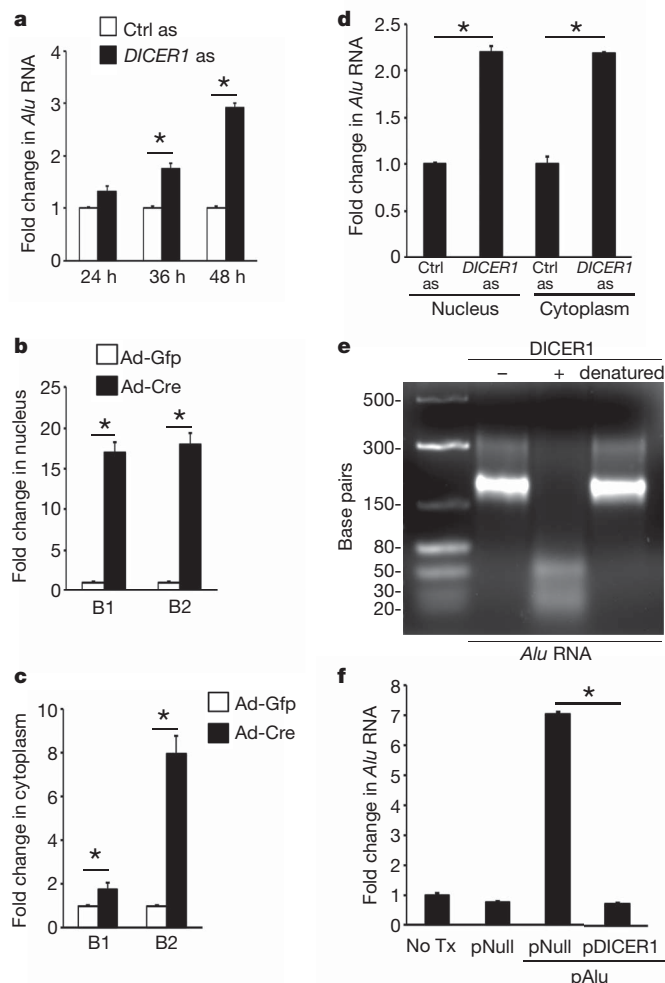


Figure 3 | *DICER1* degrades *Alu* RNA. **a**, *DICER1* antisense (as) increased *Alu* RNA in human RPE cells. **b**, **c**, Ad-Cre, but not Ad-GFP, increased B1 and B2 RNAs in *Dicer1*^{fl/fl} mouse RPE cells in nucleus (**b**) and cytoplasm (**c**). **d**, *DICER1* as upregulated *Alu* RNA in human RPE cell nucleus and cytoplasm. **e**, Agarose gel electrophoresis shows recombinant *DICER1* (+), but not heat-denatured *DICER1*, degrades *Alu* RNA isolated and cloned from human GA RPE. Image representative of six experiments. **f**, *Alu* RNA in human RPE cells upregulated by plasmid coding for *Alu* (p*Alu*) versus pNull or no treatment (no Tx) at 24 h reduced by p*DICER1*. * $P < 0.05$. $n = 4$ –8 (**a**–**d**, **f**). Values normalized to control as-treated (for *Alu*) or Ad-GFP-infected cells (for B elements).

increase in the abundance of RNAs coded by L1.3, human endogenous retrovirus-W envelope, or hY3 (Supplementary Fig. 14). These data demonstrate that *Alu* RNA accumulation is a biologically specific response to *DICER1* depletion.

Alu RNA upregulation induced by *DICER1* knockdown was inhibited by taqetoxin (an RNA polymerase III inhibitor) but not α -amanitin (an RNA polymerase II inhibitor) (Supplementary Fig. 15). Northern blotting showed that *Alu* RNA from the RPE of human eyes with GA was approximately 300 nucleotides in length, consistent with the length of non-embedded polymerase III *Alu* transcripts. Our northern probe specifically detected *Alu* RNA and not 7SL RNA, the evolutionary precursor of *Alu*. Northern blotting showed no difference in 7SL RNA abundance between the RPE of GA and control eyes. Real-time PCR with reverse transcription analysis showed that 7SL RNA was not dysregulated in the RPE of human eyes with GA or in *DICER1*-depleted human RPE cells (Supplementary Fig. 16). *DICER1* knockdown did not upregulate several polymerase II-transcribed genes (*ADAR2* (also known as *ADARB1*), *NICN*, *NLRP* and *SLFN11*) containing exon-embedded *Alu* sequences. These data indicate that *Alu* RNA in the

RPE of human eyes with GA are primary *Alu* transcripts and not passenger or bystander sequences embedded in other RNAs. Conclusive assignment of these *Alu* sequences as polymerase III transcripts must await precise determination of their transcription start site.

We tested whether *Alu* RNA accumulation could induce GA. Transfecting human or wild-type mouse RPE cells with a plasmid coding for *Alu* (p*Alu*) reduced cell viability (Supplementary Fig. 17). Subretinal transfection of plasmids coding for two different *Alu* RNAs or for B1 or B2 RNAs induced RPE degeneration in wild-type mice (Fig. 4a, Supplementary Fig. 17 and data not shown). Treating human RPE cells with a recombinant 281-nucleotide-long polymerase III-derived *Alu* RNA isolated from a human embryonal carcinoma cell line dose-dependently increased cell death (Fig. 4b), indicating that endogenous DICER1 degrades small amounts of *Alu* RNA but can be overwhelmed. Accordingly, DICER1 overexpression blocked p*Alu*-induced cell death in human RPE cells and RPE degeneration in wild-type mice (Supplementary Fig. 17).

Subretinal injection delivered *Alu* RNA to RPE cells in wild-type mice (Supplementary Fig. 18), consistent with the ability of long RNAs with duplex motifs to enter cells²⁶. We cloned a 302-nucleotide-long *Alu* RNA isolated from the RPE of a human eye with GA and transcribed it *in vitro* to generate partially and completely annealed structures that mimic *Alu* RNAs transcribed by polymerase III and polymerase II, respectively. Subretinal injection of either of these *Alu* RNAs induced RPE degeneration in wild-type mice (Fig. 4c and Supplementary Fig. 19), supporting the assignment of disease causality. In contrast, subretinal injection of these *Alu* RNAs digested with DICER1 did not induce RPE degeneration (Fig. 4d and Supplementary Fig. 19). When these *Alu* RNAs were subjected to mock DICER1 digestion, they induced RPE degeneration, suggesting a role for DICER1 in protecting against *Alu* RNA-induced degeneration.

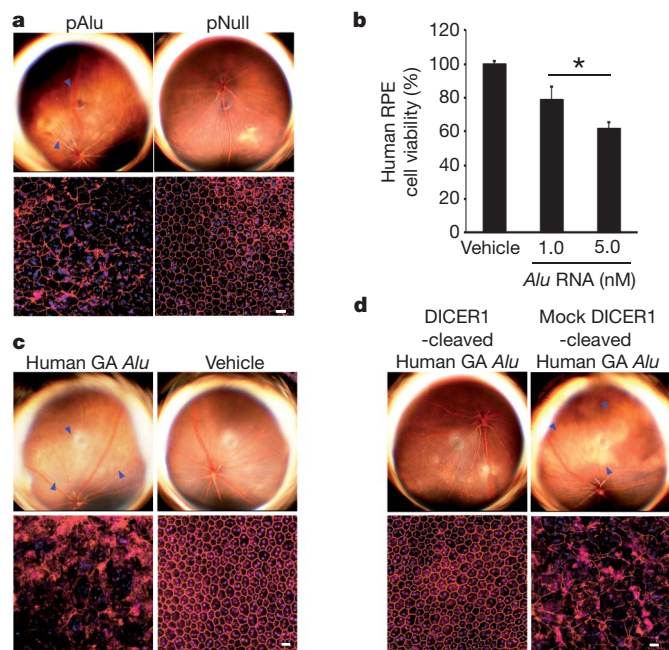


Figure 4 | *DICER1* protects RPE cells from *Alu* RNA cytotoxicity.

a, Subretinal p*Alu*, but not p*Null*, induced wild-type mouse RPE degeneration (fundus photographs, top row; ZO-1 stained (red) flat mounts, bottom row). **b**, *Alu* RNA induced human RPE cytotoxicity. Values normalized to p*Null* or vehicle. * $P < 0.05$ by Student *t*-test. $n = 4-6$. **c**, Subretinal *Alu* RNA isolated and cloned from human GA RPE induced wild-type mouse RPE degeneration. **d**, Subretinal injection of this *Alu* RNA, when cleaved by DICER1, did not induce wild-type mouse RPE degeneration (fundus photographs, top row; flat mounts, bottom row) in contrast to mock-cleaved *Alu* RNA. Degeneration outlined by blue arrowheads (**a**, **c**, **d**). Scale bars, 20 μm . $n = 10-15$.

In contrast, subretinal transfection of transfer RNA or plasmids coding for 7SL RNA or two different primary miRNAs did not induce RPE degeneration in wild-type mice (Supplementary Fig. 20). Chemically synthesized dsRNAs that mimic viral dsRNA can induce RPE degeneration by activating toll like receptor-3 (TLR3)²⁷; however, p*Alu* transfection did not induce TLR3 phosphorylation in human RPE cells and did induce RPE degeneration in *Tlr3*^{-/-} mice (Supplementary Fig. 21). Therefore, *Alu* RNA-induced RPE degeneration cannot be attributed solely to its repetitive or double-stranded nature, as it exerted effects distinct from other structured dsRNAs of similar length.

The mechanism of RPE cell death in GA is undefined. DNA fragmentation has been identified in RPE cells in human eyes with GA²⁸, and *Dicer1* knockdown has been associated with induction of apoptosis in diverse tissues^{10,29}. We now provide evidence of caspase-3 cleavage in regions of RPE degeneration in human eyes with GA (Supplementary Fig. 22). Caspase-3 cleavage was also observed in the RPE cells of *BEST1* Cre; *Dicer1*^{fl/fl} mice and in *Alu* RNA-stimulated or -overexpressing human RPE cells. These data indicate a role for *Alu* RNA-induced RPE cell apoptosis triggered by DICER1 dysregulation in GA.

To study whether an imbalance in small RNA species produced from long *Alu* RNAs could contribute to RPE degeneration, we exposed human RPE cells or wild-type mice to DICER1 cleavage fragments of *Alu* RNA. Subretinal transfection of these fragments did not damage RPE cells in wild-type mice, and co-administering these fragments did not prevent RPE cell degeneration in wild-type mice induced by p*Alu* (Supplementary Fig. 23). Similarly, these fragments did not prevent human RPE cell death induced by *Alu* RNA overexpression. These data indicate that upregulation of long *Alu* RNA rather than imbalance in *Alu* RNA-derived small RNA fragments is responsible for RPE degeneration induced by DICER1 reduction.

To dissect the contribution of *Alu* RNA accumulation versus that of miRNA dysregulation to RPE degeneration in the context of DICER1 deficit, we re-examined HCT-DICER1^{ex5} cells in which miRNA biogenesis is impaired but long dsRNA cleavage is preserved due to the intact RNase III domains. *Alu* RNA levels were not different between HCT-DICER1^{ex5} and parent HCT116 cells (Supplementary Fig. 24). In contrast, DICER1 knockdown in HCT116 cells upregulated *Alu* RNA. Also, *Alu* RNA induces similar levels of cytotoxicity in HCT-DICER1^{ex5} and parent HCT116 cells, indicating that coexisting miRNA expression deficits do not augment *Alu* RNA-induced RPE degeneration. In conjunction with the discordance in the RPE degeneration phenotype between ablation of *Dicer1* and that of various other small RNA biogenesis pathway genes in mice, our findings indicate that *Alu* RNA accumulation is critical to DICER1 reduction-induced cytotoxicity.

RPE degeneration blocked by *Alu* RNA inhibition

We tested whether DICER1 reduction-induced cytotoxicity is due to *Alu* RNA accumulation. *DICER1*-knockdown-induced human RPE cytotoxicity was inhibited by antisense oligonucleotides targeting *Alu* RNA sequences, but not by scrambled antisense control (Fig. 5a and Supplementary Fig. 25). Ad-Cre infection of *Dicer1*^{fl/fl} mouse RPE cells reduced cell viability, and this was blocked by antisense oligonucleotides targeting B1/B2 RNAs but not by scrambled antisense control (Fig. 5b and Supplementary Fig. 25). Subretinal administration of antisense oligonucleotides that reduced accumulation of B1/B2 RNAs inhibited RPE degeneration in AAV1-*BEST1*-Cre-treated *Dicer1*^{fl/fl} mice (Fig. 5c and Supplementary Fig. 25), providing evidence of *in vivo* functional rescue.

We tested whether *Alu* inhibition also rescued miRNA expression deficits as a potential explanation for the functional rescue of DICER1 depletion-induced RPE degeneration. As expected, *DICER1* knockdown in human RPE cells reduced the abundance of numerous miRNAs (Fig. 5d). However, inhibition of *Alu* RNA did not recover miRNA expression. Thus, rescue of RPE cell viability by *Alu* RNA inhibition despite persistent global miRNA expression deficits shows

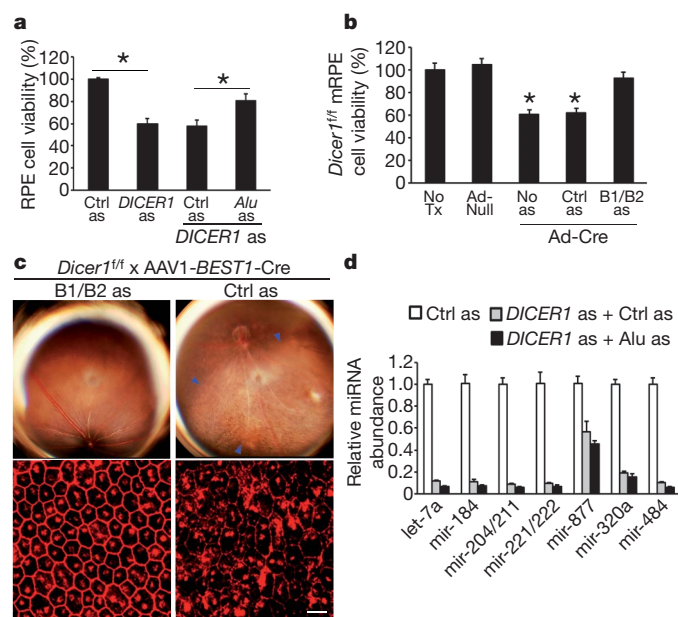


Figure 5 | *DICER1* dysregulation induces RPE cell death via *Alu* RNA accumulation. **a**, Human RPE cytotoxicity induced by *DICER1* as rescued by *Alu* RNA antisense. Values normalized or compared to control (Ctrl) antisense. **b**, Ad-Cre but not Ad-Null induced *Dicer1^{fl}* mouse RPE cytotoxicity. B1/B2 RNA as, but not control (Ctrl) as, rescued viability. Values normalized to untreated cells (no Tx). * $P < 0.05$ by Student *t*-test. $n = 4-6$ (**a**, **b**). **c**, Subretinal AAV-BEST1-Cre induced RPE degeneration (blue arrowheads in fundus photograph, top row; ZO-1 stained (red) flat mounts, bottom row) in *Dicer1^{fl}* mice 20 days after injection was inhibited by subretinal cholesterol-conjugated B1/B2 as, but not cholesterol-conjugated Ctrl as, 10 days after AAV-BEST1-Cre injection. Values normalized to Ctrl as-treatment. $n = 8$. Scale bar, 20 μ m. **d**, *DICER1* as induced global miRNA expression deficits in human RPE cells compared to Ctrl as. No significant difference in miRNA abundance between *Alu* as and Ctrl as-treated *DICER1* depleted cells. $n = 3$.

that RPE degeneration induced by *DICER1* deficit is due to *Alu* RNA accumulation and not miRNA dysregulation.

Collectively, these data support a model in which primary *Alu* transcripts are responsible for RPE degeneration. Whether similar pathology can also result from upregulation of as yet undefined polymerase II transcripts with embedded *Alu* sequences is an intriguing possibility that requires further study. Importantly, we demonstrated that primary *Alu* transcripts are elevated in human disease, that *Alu* transcripts recapitulate disease in relevant experimental models, and that targeted suppression of *Alu* transcripts successfully inhibits this pathology. These observations have direct relevance for clinical strategies to prevent and treat GA.

Discussion

Our findings elucidate a critical cell survival function for *DICER1* by functional silencing of toxic *Alu* transcripts. This unexpected function suggests that RNAi-independent mechanisms should be considered in interpreting the phenotypes of systems in which *DICER1* is dysregulated. For example, it would be interesting to test whether *DICER1* ablation induced cytotoxicity in mouse neural retina³⁰ and heart³¹ might also involve B1/B2 RNA accumulation. More broadly, recognition of *DICER1*'s hitherto unidentified function as an important controller of transcripts derived from the most abundant genomic repetitive elements can illuminate new functions for RNases in cytoprotective surveillance. *DICER1* expression is reduced in GA and partial loss of *DICER1* promotes RPE degeneration; thus loss of heterozygosity in *DICER1* may underlie the development of GA, similar to its function as a haploinsufficient tumour suppressor³²⁻³⁴.

This also is, to our knowledge, the first example of how *Alu* could cause a human disease via direct RNA cytotoxicity rather than by

inducing chromosomal DNA rearrangements or insertional mutagenesis through retrotransposition, which have been implicated in diseases such as α -thalassaemia³⁵, colon cancer³⁶, hypercholesterolemia^{37,38}, and neurofibromatosis³⁹. Future studies should determine the precise chromosomal locus of the *Alu* RNA elements that accumulate in GA and the nature of transcriptional and post-transcriptional machinery that enable their biogenesis.

In addition to processing miRNAs⁵, *DICER1* has been implicated in heterochromatin assembly^{40,41}. Since *Alu* elements are abundant within heterochromatin⁴², whether perturbations in centromeric silencing underlie the pathogenesis of GA warrants study. The finding that chromatin remodelling at *Alu* repeats can regulate miRNA expression⁴³ raises the intriguing possibility of other regulatory intersections between *DICER1* and *Alu*. It also remains to be investigated whether centromeric satellite repeats that accumulate in *Dicer1*-null mouse embryonic stem cells^{44,45} might be involved in the pathogenesis of GA.

In the mouse germline, *Dicer1* has been implicated in generating endogenous small interfering RNAs (endo-siRNAs) from repeat elements^{46,47}. If this process is conserved in mammalian somatic tissues, it would be interesting to learn whether endo-siRNAs serve a homeostatic function in preventing the development of GA. Given that caspases can cleave *Dicer1* and convert it into a DNase that promotes apoptosis in nematodes⁴⁸, our finding that *Alu* RNA induces caspase activation introduces the possibility of bidirectional regulation between *DICER1* and *Alu* that triggers feed-forward disease-amplifying loops.

The inciting events that trigger an RPE-specific reduction of *DICER1* in patients with GA are unknown. Potential culprits could include oxidative stress, which is postulated to underlie AMD pathogenesis⁴, as we found that hydrogen peroxide downregulates *DICER1* in human RPE cells (Supplementary Fig. 26). Whereas upstream triggers of *DICER1* dysregulation and the role of other *DICER1*-dependent, DROSHA/DGCR8-independent small RNAs in GA await clarification, the ability of *Alu* RNA antisense oligonucleotides to inhibit RPE cytotoxicity induced by *DICER1* depletion provides a rationale to investigate *Alu* RNA inhibition or *DICER1* augmentation as potential therapies for GA.

METHODS SUMMARY

Subretinal injections (1 μ l) were performed using a Pico-Injector (PLI-100, Harvard Apparatus). Plasmids were transfected *in vivo* using 10% Neuroporter (Genlantis). Immunolabelling was performed using antibodies against dsRNA (clone J2, English & Scientific Consulting), *DICER1* (Santa Cruz Biotechnology), zonula occludens-1 (Invitrogen), Cre recombinase (EMD4Biosciences), or cleaved caspase-3 (Cell Signaling). dsRNA was isolated by immunoprecipitating homogenized tissue lysates with 40 μ g of J2 for 16 h at 4 °C. Purified dsRNA was ligated to an anchor primer and purified by MinElute Gel extraction columns (Qiagen). Ligated dsRNA was denatured, reverse transcribed, and amplified by PCR. Amplified cDNA products were cloned into pCRII TOPO vector (Invitrogen) and sequenced. Homology to *Alu* consensus sequences was determined using CENSOR. Cell viability was assessed using CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega). Total RNA (1 μ g) was reverse transcribed using qScript cDNA SuperMix (Quanta Biosciences) and amplified by real-time quantitative PCR (Applied Biosystems 7900 HT) with Power SYBR green Master Mix. Relative expressions were determined by the $2^{-\Delta\Delta Ct}$ method. miRNA abundance was quantified using All-in-One miRNA qRT-PCR Detection Kit (GeneCopoeia).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 July 2010; accepted 18 January; corrected 17 March 2011 (see full-text HTML version for details).

Published online 6 February 2011.

- Kleinman, M. E. *et al.* Sequence- and target-independent angiogenesis suppression by siRNA via TLR3. *Nature* **452**, 591–597 (2008).
- Takeda, A. *et al.* CCR3 is a target for age-related macular degeneration diagnosis and therapy. *Nature* **460**, 225–230 (2009).
- Ferrara, N. Vascular endothelial growth factor and age-related macular degeneration: from basic science to therapy. *Nature Med.* **16**, 1107–1111 (2010).

4. Ambati, J., Ambati, B. K., Yoo, S. H., Ianchulev, S. & Adamis, A. P. Age-related macular degeneration: etiology, pathogenesis, and therapeutic strategies. *Surv. Ophthalmol.* **48**, 257–293 (2003).
5. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
6. Batzer, M. A. & Deininger, P. L. Alu repeats and human genomic diversity. *Nature Rev. Genet.* **3**, 370–379 (2002).
7. Gregory, R. I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235–240 (2004).
8. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–1441 (2004).
9. Meister, G. *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15**, 185–197 (2004).
10. Harfe, B. D., McManus, M. T., Mansfield, J. H., Hornstein, E. & Tabin, C. J. The RNaseIII enzyme *Dicer* is required for morphogenesis but not patterning of the vertebrate limb. *Proc. Natl Acad. Sci. USA* **102**, 10898–10903 (2005).
11. Iacovelli, J. *et al.* Generation of cre transgenic mice with postnatal RPE-specific ocular expression. *Invest. Ophthalmol. Vis. Sci.* doi:10.1167/iov.10-6347 (6 January 2011).
12. Alexander, J. J. & Hauswirth, W. W. Adeno-associated viral vectors and the retina. *Adv. Exp. Med. Biol.* **613**, 121–128 (2008).
13. Chong, M. M., Rasmussen, J. P., Rudensky, A. Y. & Littman, D. R. The RNaseIII enzyme *Drosha* is critical in T cells for preventing lethal inflammatory disease. *J. Exp. Med.* **205**, 2005–2017 (2008).
14. Yi, R. *et al.* DGCR8-dependent microRNA biogenesis is essential for skin development. *Proc. Natl Acad. Sci. USA* **106**, 498–502 (2009).
15. O'Carroll, D. *et al.* A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. *Genes Dev.* **21**, 1999–2004 (2007).
16. Chong, M. M. *et al.* Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev.* **24**, 1951–1960 (2010).
17. Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P. & Belloch, R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, *Dicer*-dependent small RNAs. *Genes Dev.* **22**, 2773–2785 (2008).
18. Schaefer, A. *et al.* Argonaute 2 in dopamine 2 receptor-expressing neurons regulates cocaine addiction. *J. Exp. Med.* **207**, 1843–1851 (2010).
19. Diederichs, S. & Haber, D. A. Dual role for Argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **131**, 1097–1108 (2007).
20. Kaneda, M., Tang, F., O'Carroll, D., Lao, K. & Surani, M. A. Essential role for Argonaute2 protein in mouse oogenesis. *Epigenetics Chromatin* **2**, 9 (2009).
21. Su, H., Trombly, M. I., Chen, J. & Wang, X. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes Dev.* **23**, 304–317 (2009).
22. Chendrimada, T. P. *et al.* TRBP recruits the *Dicer* complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740–744 (2005).
23. Cummins, J. M. *et al.* The colorectal microRNAome. *Proc. Natl Acad. Sci. USA* **103**, 3687–3692 (2006).
24. Schonborn, J. *et al.* Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. *Nucleic Acids Res.* **19**, 2993–3000 (1991).
25. Kato, H. *et al.* Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-1 and melanoma differentiation-associated gene 5. *J. Exp. Med.* **205**, 1601–1610 (2008).
26. Saleh, M. C. *et al.* The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. *Nature Cell Biol.* **8**, 793–802 (2006).
27. Yang, Z. *et al.* Toll-like receptor 3 and geographic atrophy in age-related macular degeneration. *N. Engl. J. Med.* **359**, 1456–1463 (2008).
28. Dunaief, J. L., Dentchev, T., Ying, G. S. & Milam, A. H. The role of apoptosis in age-related macular degeneration. *Arch. Ophthalmol.* **120**, 1435–1442 (2002).
29. Davis, T. H. *et al.* Conditional loss of *Dicer* disrupts cellular and tissue morphogenesis in the cortex and hippocampus. *J. Neurosci.* **28**, 4322–4330 (2008).
30. Damiani, D. *et al.* *Dicer* inactivation leads to progressive functional and structural degeneration of the mouse retina. *J. Neurosci.* **28**, 4878–4887 (2008).
31. Chen, J. F. *et al.* Targeted deletion of *Dicer* in the heart leads to dilated cardiomyopathy and heart failure. *Proc. Natl Acad. Sci. USA* **105**, 2111–2116 (2008).
32. Merritt, W. M. *et al.* *Dicer*, *Drosha*, and outcomes in patients with ovarian cancer. *N. Engl. J. Med.* **359**, 2641–2650 (2008).
33. Kumar, M. S. *et al.* *Dicer1* functions as a haploinsufficient tumor suppressor. *Genes Dev.* **23**, 2700–2704 (2009).
34. Hill, D. A. *et al.* *DICER1* mutations in familial pleuropulmonary blastoma. *Science* **325**, 965 (2009).
35. Nicholls, R. D., Fischel-Ghodsian, N. & Higgs, D. R. Recombination at the human α -globin gene cluster: sequence features and topological constraints. *Cell* **49**, 369–378 (1987).
36. Nyström-Lahti, M. *et al.* Founding mutations and Alu-mediated recombination in hereditary colon cancer. *Nature Med.* **1**, 1203–1206 (1995).
37. Lehrman, M. A. *et al.* Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* **227**, 140–146 (1985).
38. Lehrman, M. A., Goldstein, J. L., Russell, D. W. & Brown, M. S. Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell* **48**, 827–835 (1987).
39. Wallace, M. R. *et al.* A *de novo* Alu insertion results in neurofibromatosis type 1. *Nature* **353**, 864–866 (1991).
40. Volpe, T. A. *et al.* Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833–1837 (2002).
41. Hall, I. M. *et al.* Establishment and maintenance of a heterochromatin domain. *Science* **297**, 2232–2237 (2002).
42. Prades, C., Laurent, A. M., Puechberty, J., Yurov, Y. & Roizes, G. SINE and LINE within human centromeres. *J. Mol. Evol.* **42**, 37–43 (1996).
43. Saito, Y. *et al.* Chromatin remodeling at Alu repeats by epigenetic treatment activates silenced *microRNA-512-5p* with downregulation of *Mcl-1* in human gastric cancer cells. *Oncogene* **28**, 2738–2744 (2009).
44. Murchison, E. P., Partridge, J. F., Tam, O. H., Cheloufi, S. & Hannon, G. J. Characterization of *Dicer*-deficient murine embryonic stem cells. *Proc. Natl Acad. Sci. USA* **102**, 12135–12140 (2005).
45. Kanellopoulou, C. *et al.* *Dicer*-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* **19**, 489–501 (2005).
46. Tam, O. H. *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**, 534–538 (2008).
47. Watanabe, T. *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539–543 (2008).
48. Nakagawa, A., Shi, Y., Kage-Nakadai, E., Mitani, S. & Xue, D. Caspase-dependent conversion of *Dicer* ribonuclease into a death-promoting deoxyribonuclease. *Science* **328**, 327–334 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Chrenek, J. Garcia-Perez, T. Heidmann, C. Kanellopoulou, D. M. Livingston, J. V. Moran, R. F. Mullins, J. M. Nickerson, E. A. Pearce, A. Tarakhovskiy, B. Vogelstein, V. E. Velculescu and D. J. Zack for providing mice, reagents or tissues; R. King, L. Xu, M. McConnell, C. Payne, G. R. Pattison, G. J. Jaffe, S. Medearis and C. Spee for technical assistance; and A. Sinai, R. Mohan, T. S. Khurana, R. A. Brekken, P. L. Deininger, S. Bondada, P. A. Pearson, A. M. Rao, G. S. Rao and K. Ambati for discussions. J.A. was supported by National Eye Institute (NEI)/National Institutes of Health (NIH) grants R01EY015422, R01EY018350, R01EY018836, R01EY020672, R21EY019778, RC1EY020442, the Doris Duke Distinguished Clinical Scientist Award, the Burroughs Wellcome Fund Clinical Scientist Award in Translational Research, and the Dr E. Vernon Smith and Eloise C. Smith Macular Degeneration Endowed Chair. Research to Prevent Blindness Senior Scientist Investigator Awards or departmental unrestricted grants supported J.A., H.E.G. and W.W.H.; J.Z.B. was supported by University of Kentucky Physician Scientist Award, International Retinal Research Foundation, and American Health Assistance Foundation; B.K.A. by VA Merit Award and Department of Defense; D.-k.L. by Global Research Laboratory program by MEST, Korea; D.R.H. by Arnold and Mabel Beckman Foundation; W.W.H. by Macular Vision Research Foundation and Foundation Fighting Blindness; J.M.P. by ARC Centres of Excellence Grant CE0561903; M.C.M. by Sydney Foundation for Medical Research. B.K.A. was supported by NIH R01EY017182 and R01EY017950; R.E.B. by NIH R01HD027215; G.C. by NIH R21AI076757; H.E.G. by NIH P30EY06360; W.W.H. by NIH U10EY013729, R01EY011123, and P30EY008571; J.F.K. and J.A.G. by NIH R01GM068414; J.L.D. by NIH R01EY015240; D.R.H. by NIH P30EY003040 and R01EY001545; M.E.K. and S.B. by NIH T32HL091812. P.P. is a Senior Scholar from the Fonds de la Recherche en Santé du Québec (FRSQ). M.M.W.C. is a QEII Fellow of the Australian Research Council and is supported by National Health and Medical Research Council, Australia Project Grant 637228. E.F. and D.R.L. are investigators of the Howard Hughes Medical Institute.

Author Contributions H.K., S.D., V.T., W.G.C., B.J.F., M.E.K., S.L.P., J.F.K., J.A.G., K.K., N.L.J., B.D.G., Y.H., R.J.C.A., A.D.B., S.B., J.W., M.H., Y.S. and J.Z.B. performed experiments. W.W.H., V.A.C., D.-k.L., J.W.Y., C.M.R., D.R.H., H.E.G., Q.Z., J.M.P., M.C.M., A.H.M., M.M.W.C., D.R.L., E.F., P.P., F.W.B., R.E.B., S.M., G.C. and J.L.D. provided tissues or reagents. J.A. conceived and directed the project, and wrote the paper with assistance from P.P., C.M.R., K.K., J.F.K., J.A.G., E.F., M.M.W.C., B.J.F., B.D.G. and B.K.A. All authors had the opportunity to discuss the results and comment on the manuscript.

Author Information The *Alu* sequences have been deposited in GenBank under the accession numbers HN176584 and HN176585. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.A. (jamba2@email.uky.edu).

METHODS

Human tissue. Donor eyes or ocular tissues from patients with GA due to AMD or patients without AMD were obtained from various eye banks in Australia and the United States of America. These diagnoses were confirmed by dilated ophthalmic examination before acquisition of the tissues or eyes or upon examination of the eye globes post mortem. The study followed the guidelines of the Declaration of Helsinki. Institutional review boards granted approval for allocation and histological analysis of specimens.

Animals. All animal experiments were in accordance with the guidelines of the University of Kentucky Institutional Animal Care and Use Committee and the Association for Research in Vision and Ophthalmology.

Immunolabelling and histology. Fixed human tissue was stained with the antibodies against dsRNA (clone J2, English & Scientific Consulting) or human DICER1 (Santa Cruz Biotechnology). Bound antibody was detected with alkaline phosphatase streptavidin solution (Invitrogen) and the enzyme complex was visualized by Vector Blue (Vector Laboratories). Mouse RPE/choroid flat mounts were fixed with 4% paraformaldehyde or 100% methanol and stained with rabbit antibodies against human zonula occludens-1 (Invitrogen), Cre recombinase (EMD4Biosciences), or human cleaved caspase-3 (Cell Signaling) and visualized with Alexa594- or Cy5-conjugated secondary antibodies. Fixed primary human RPE cells were stained with antibodies against dsRNA or human DICER1 and visualized with Alexa Fluor-conjugated secondary antibodies. Nuclei were visualized with DAPI counterstaining.

Subretinal injections. Subretinal injections (1 μ l) in mice were performed using a Pico-Injector (PLI-100, Harvard Apparatus). *In vivo* transfection of plasmids was achieved using 10% Neuroporter (Genlantis). AAV1-BEST1-Cre¹² or AAV1-BEST1-GFP were injected at 1.0×10^{11} p.f.u. ml⁻¹ and recombinant Alu RNAs were injected at 0.3 mg ml⁻¹. Cell-permeating cholesterol-conjugated B1/B2 antisense oligonucleotides or cholesterol-conjugated control antisense (both from Integrated DNA Technologies) were injected (2 μ g in 1 μ l) 10 days after AAV1-BEST1-Cre was injected in Dicer1^{fl/fl} mice.

dsRNA isolation. Human eyes were stored in RNAlater (Ambion). Homogenized lysates were immunoprecipitated with 40 μ g of mouse antibody against dsRNA (clone J2) for 16 h at 4 °C. Immunocomplexes were collected on protein A/G agarose (Thermoscientific) and dsRNA species were separated and isolated using Trizol (Invitrogen) according to the manufacturer's instructions. Purified dsRNA was then ligated to an anchor primer and purified by MinElute Gel extraction columns (Qiagen). Ligated dsRNA was then denatured, reverse transcribed into cDNA, and amplified by PCR. Amplified cDNA products were cloned into PCRII TOPO vector (Invitrogen) and sequenced at the University of Kentucky Advanced Genetic Technologies Center. The homology of isolated cDNA sequences to known Alu consensus sequences was determined using the CENSOR server⁴⁹.

Cell culture. All cell lines were cultured at 37 °C and 5% CO₂. Primary mouse RPE cells were isolated as previously described⁵⁰ and grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS and standard antibiotics concentrations. Primary human RPE cells were isolated as previously described²⁷ and maintained in DMEM supplemented with 20% FBS and antibiotics. Parental HCT116 and isogenic Dicerex5 cells²³ were grown in McCoy's 5A medium supplemented with 10% FBS. Transient transfections of plasmid and antisense oligonucleotides were performed with Lipofectamine2000 (Invitrogen) and Oligofectamine (Invitrogen), respectively. Cell viability measurements were performed using the CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega) according to the manufacturer's instructions.

Alu RNA synthesis. Two Alu RNAs were synthesized: a 281-nucleotide Alu sequence originating from the cDNA clone TS 103 which is known to be expressed in human cells⁵¹, and a 302-nucleotide Alu sequence isolated from the RPE of a human eye with GA. Alu RNAs were synthesized using a RNA polymerase T7 promoter and runoff transcription followed by gel purification as previously described⁵², yielding ssRNAs that fold into a defined secondary structure identical to polymerase III derived transcripts. We also synthesized a fully complementary dsRNA form (resembling a polymerase II derived transcript) of the 302-nucleotide human GA Alu using linearized PCRII TOPO plasmid templates using T7 or SP6 RNA polymerases (MegaScript, Ambion) according to the manufacturer's recommendations. After purification, equal molar amount of each transcript were combined and heated at 95 °C for 1 min, cooled and then annealed at room temperature for 24 h. The Alu dsRNA was precipitated, suspended in water and analysed on 1.4% non-denaturing agarose gel using the single-stranded complementary strands as controls.

Real-time PCR. Total RNA was extracted from tissues or cells using TRIzol reagent (Invitrogen) according to manufacturer's recommendations and were treated with RNase free DNase (Ambion). Total RNA (1 μ g) was reverse transcribed as previously described² using qScript cDNA SuperMix (Quanta Biosciences). The RT products (cDNA) were amplified by real-time quantitative

PCR (Applied Biosystems 7900 HT Fast Real-Time PCR system) with Power SYBR green Master Mix. Oligonucleotide primers specific for *DICER1* (forward 5'-CCCGGCTGAGAGAACTTACG-3' and reverse 5'-CTGTAACCTCGACCAACACCTTTAAA-3'), *DROSHA* (forward 5'-GAACAGTTCAACCCCGATG TG-3' and reverse 5'-CTCAACTGTGCAGGGCGTATC-3'), *DGCR8* (forward 5'-TCTGCTCCTTAGCCCTGTCACT-3' and reverse 5'-CCAACACTCCCGC CAAAG-3'), *EIF2C2* (forward 5'-GCACGGAAGTCCATCTGAAGTC-3' and reverse 5'-CCGGCGTCTCTCGAGATCT-3'), human 18S rRNA (forward 5'-CGCAGCTAGGAATAATGGAATAGG-3' and reverse 5'-GCCTCAGTTCC GAAAACCAA-3'), *Alu* (forward 5'-CAACATAGTGAAACCCCGTCTCT-3' and reverse 5'-GCCTCAGCTCCCGAGTAG-3'), LINE L1.3 (ORF2) (forward 5'-CGGTGATTTCTGCATTTCCTCA-3' and reverse 5'-TGCTGGCACTCCCT AGTGAGA-3'), *HERV-WE1* (forward 5'-GCCGCTGTATGACCAAGTAGCT-3' and reverse 5'-GGGACGTCGATTCTCCAT-3'), human Ro-associated Y3 (*hY3*) (forward 5'-CCGAGTGCAGTGGTGTTCACA-3' and reverse 5'-GGA GTGGAGAAGGAACAAAGAAATC-3'), 7SL (forward 5'-CGGCATCAA TATGGTGACCT-3' and reverse 5'-CTGATCAGCAGCGGAGTTC-3'), B1 (forward 5'-TGCCTTTAATCCCGACACTT-3' and reverse 5'-GCTGCTCA CACAAGGTTGAA-3'), B2 (forward 5'-GAGTTCAAATCCCGACCAACCA-3' and reverse 5'-AAGAGGGTCTCAGATCTTGTACAGA-3'), *Dicer1* (forward 5'-CCCACCGAGGTGCATGTT-3' and reverse 5'-TAGTGGTAGGAGGCGTG TGTAATAA-3'), mouse 18S rRNA (forward 5'-TTCGATTGCGCCGCTAGA-3' and reverse 5'-CTTTCGCTCTGCTGCCGCTCT-3') were used. The qPCR cycling conditions were 50 °C for 2 min, 95 °C for 10 min followed by 40 cycles of a two-step amplification program (95 °C for 15 s and 58 °C for 1 min). At the end of the amplification, melting curve analysis was applied using the dissociation protocol from the Sequence Detection system to exclude contamination with unspecific PCR products. The PCR products were also confirmed by agarose gel and showed only one specific band of the predicted size. For negative controls, no RT products were used as templates in the QPCR and verified by the absence of gel-detected bands. Relative expressions of target genes were determined by the 2^{- $\Delta\Delta C_t$} method.

miRNA PCR. miRNA abundance was quantified using the All-in-One miRNA qRT-PCR Detection Kit (GeneCopoeia). Briefly, total RNA was polyadenylated and reverse transcribed using a poly dT-adaptor primer. Quantitative RT-PCR was carried out using a miRNA-specific forward primer and universal reverse primer. PCR products were subjected to dissociation curve and gel electrophoresis analyses to ensure that single, mature miRNA products were amplified. Data were normalized to *ACTB* levels. The forward primers for the miRNAs were as follows: miR-184 (5'-TGGACGAGAACTGATAAGGGT-3'); miR-221/222 (5'-AGC TACATCTGGCTACTGGGT-3'); miR-204/211 (5'-TTCCCTTTGTCATCCT TCGCCT-3'); miR-877 (5'-GTAGAGGAGATGGCGCAGGG-3'); miR-320a (5'-AAAAGCTGGGTTGAGAGGGCGA-3'); miR-484 (5'-TCAGGCTCAGTC CCCTCCCGAT-3'); let-7a (5'-TGAGGTAGTAGTTGTATAGTT-3'). The reverse primers were proprietary (GeneCopoeia). The primers for *ACTB* were forward (5'-TGGATCAGCAAGCAGGAGTATG-3') and reverse (5'-GCATT TGCGGTGGACGAT-3').

Western blot. Tissues were homogenized in lysis buffer (10 mM Tris base, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.5% NP-40, protease and phosphatase inhibitor cocktail (Roche)). Protein concentrations were determined using a Bradford assay kit (Bio-Rad) with bovine serum albumin as a standard. Proteins (40–100 μ g) were run on 4–12% Novex Bis-Tris gels (Invitrogen). The transferred membranes were blocked for 1 h at room temperature and incubated with antibodies against DICER1 (1:1,000, ref. 45; or 1:200, Santa Cruz Biotechnology) at 4 °C overnight. Protein loading was assessed by immunoblotting using an anti-Tubulin antibody (1:1,000; Sigma-Aldrich). The secondary antibodies were used (1:5,000) for 1 h at room temperature. The signal was visualized by enhanced chemiluminescence (ECL Plus) and captured by VisionWorksLS Image Acquisition and Analysis software (Version 6.7.2, UVP, LLC). Densitometry analysis was performed using ImageJ (NIH). The value of 1 was arbitrarily assigned for normal eye samples.

RNA polymerase inhibition. Human RPE cells were transfected with DICER1 or control antisense oligonucleotides using Lipofectamine 2000. After a change of medium at 6 h, the cells were incubated with 45 μ M tagetitoxin (Epicentre Technologies, Tagetin) or 10 μ g ml⁻¹ α -amanitin (Sigma-Aldrich) and the total RNA was collected after 24 h.

Cell viability. MTS assays were performed using the CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega) according to the manufacturer's instructions.

49. Kohany, O., Gentles, A. J., Hankus, L. & Jurka, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* **7**, 474 (2006).

50. Yang, P., Tyrrell, J., Han, I. & Jaffe, G. J. Expression and modulation of RPE cell membrane complement regulatory proteins. *Invest. Ophthalmol. Vis. Sci.* **50**, 3473–3481 (2009).

51. Shaikh, T. H., Roy, A. M., Kim, J., Batzer, M. A. & Deininger, P. L. cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts. *J. Mol. Biol.* **271**, 222–234 (1997).
52. Allen, T. A., Von Kaenel, S., Goodrich, J. A. & Kugel, J. F. The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nature Struct. Mol. Biol.* **11**, 816–821 (2004).

Structure and mechanism of the hexameric MecA–ClpC molecular machine

Feng Wang^{1*}, Ziqing Mei^{1*}, Yutao Qi¹, Chuangye Yan¹, Qi Hu¹, Jiawei Wang¹ & Yigong Shi¹

Regulated proteolysis by ATP-dependent proteases is universal in all living cells. Bacterial ClpC, a member of the Clp/Hsp100 family of AAA+ proteins (ATPases associated with diverse cellular activities) with two nucleotide-binding domains (D1 and D2), requires the adaptor protein MecA for activation and substrate targeting. The activated, hexameric MecA–ClpC molecular machine harnesses the energy of ATP binding and hydrolysis to unfold specific substrate proteins and translocate the unfolded polypeptide to the ClpP protease for degradation. Here we report three related crystal structures: a heterodimer between MecA and the amino domain of ClpC, a heterododecamer between MecA and D2-deleted ClpC, and a hexameric complex between MecA and full-length ClpC. In conjunction with biochemical analyses, these structures reveal the organizational principles behind the hexameric MecA–ClpC complex, explain the molecular mechanisms for MecA-mediated ClpC activation and provide mechanistic insights into the function of the MecA–ClpC molecular machine. These findings have implications for related Clp/Hsp100 molecular machines.

ATP-dependent proteases are exemplified by the bacterial complexes ClpAP, ClpCP and ClpXP, and the eukaryotic and prokaryotic proteasomes^{1–6}. Each of these degradation machines consists of an ATP-dependent, protein-unfolding unit, namely ClpA, ClpC, ClpX or the regulatory particle of the proteasome, and a protease unit, such as ClpP or the core particle of the proteasome. ClpA, ClpC and ClpX belong to the Clp/Hsp100 family of AAA+ proteins, which also include the molecular chaperones ClpB and Hsp104 (the orthologue of ClpB in yeast). ClpA, ClpC and ClpX recognize and unfold specific substrate proteins and translocate the unfolded polypeptide to ClpP for degradation, whereas ClpB and Hsp104 use the energy of ATP binding and hydrolysis to rescue misfolded and aggregated proteins^{7–10}. ClpX belongs to class 2 of the Clp/Hsp100 family, bearing one nucleotide-binding domain, whereas ClpA, ClpB, ClpC and Hsp104 belong to class 1, each containing two nucleotide-binding domains, known as D1 and D2.

The functional form of the Clp/Hsp100 proteins is hexameric. At present, no atomic structure is known for any oligomeric complex of class 1 Clp/Hsp100 proteins, which restricts mechanistic understanding of the ClpAP, ClpCP, ClpB and Hsp104 complexes. This picture is further confounded by contrasting claims about the overall architecture of the functional complexes. ClpB and Hsp104, which have considerable sequence similarity, have been proposed to adopt different hexameric structures, with their characteristic M domain placed on either the exterior^{11–13} or the interior^{14,15} of the complex. These two incompatible models lead to different interpretations^{9,16} for biochemical observations and await independent clarification.

Similar to ClpB and Hsp104, ClpC also contains an M domain. In contrast to ClpA and ClpB/Hsp104, formation of the functional ClpC complex or assembly of the ClpCP degradation machine requires the adaptor protein MecA¹⁷ (Supplementary Fig. 1a). MecA is also responsible for recruitment of specific substrate proteins such as ComK^{18,19}. Paradoxically, MecA is degraded along with ComK²⁰, resulting in the disassembly of the ClpCP protease²¹. The molecular mechanisms

behind these observations remain largely enigmatic. In particular, how MecA facilitates formation of the functional ClpC complex and the ClpCP degradation machine remains unknown.

Our current study demonstrates how MecA controls the assembly and activation of the ClpC hexamer, reveals the organizational principles of the hexameric Clp/Hsp100 proteins, resolves the controversy about the hexameric models of ClpB and Hsp104, and suggests how the MecA–ClpC molecular machine may recognize and unfold substrate proteins.

Dissection of the MecA–ClpC interactions

A carboxy-terminal domain of MecA (residues 121–218, or MecA121) was recently found to target a variety of fusion proteins to the ClpCP protease for degradation²¹. MecA121 facilitates hexamerization of ClpC and allows the activated ClpC complex to associate stably with ClpP²¹. Both the N-domain (residues 1–150) and the D1–M portion (residues 150–485) of ClpC formed a stable complex with MecA121 (Supplementary Figs 1b and 2). By contrast, D2 failed to interact with MecA121. These observations are consistent with a published report¹⁷ and suggest that MecA121 may use different surface areas to interact with the N domain and the D1–M portion of ClpC.

Next we determined the crystal structures of the ClpC N domain in isolation and in complex with MecA121 (Fig. 1a, Supplementary Table 1 and Supplementary Fig. 3a, b). The structure of the ClpC N domain, which comprises eight α -helices, is similar to that of the ClpA or ClpB N domain^{11,22,23} (Supplementary Fig. 3c, d). In the complex, helix α 2 of MecA121 stacks against three α -helices of ClpC (Fig. 1a). This interface contains seven intermolecular hydrogen bonds and a number of van der Waals contacts (Fig. 1b, c). Notably, Glu 184 of MecA, located at the C terminus of α 2, accepts three intermolecular hydrogen bonds from Thr 31 and Thr 81. Asn 177 of MecA, at the N terminus of α 2, forms two hydrogen bonds with Arg 122 and Asn 126.

Glu 184, invariant among MecA homologues, seems to have a central role in complex formation. The mutation Glu 184 Arg in

¹Center for Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China.

*These authors contributed equally to this work.

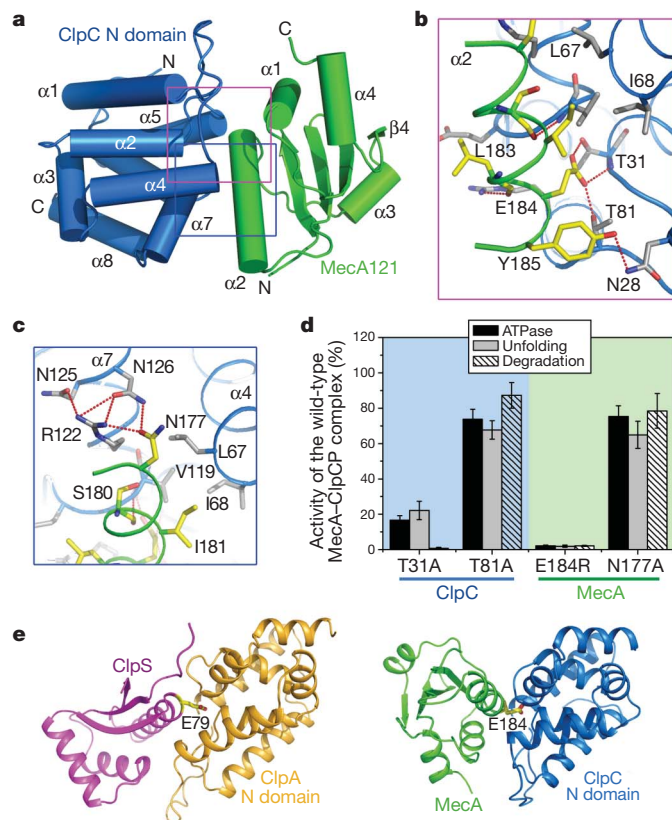


Figure 1 | A binary complex between Meca121 and the ClpC N domain. **a**, Overall structure of the complex between Meca121 (green) and ClpC N domain (blue). Interactions identified by magenta and blue boxes are detailed in **b** and **c**. All structural figures were prepared using PYMOL⁴³. **b**, Close-up view of the interactions around Glu 184. Hydrogen bonds are represented by red, dashed lines in all figures. **c**, Close-up view of the interactions around Asn 177 of Meca. **d**, Results of activity assays on the ClpC and Meca mutants. Error bars, s.d. of three sets of experiments. **e**, The Meca-ClpC N-domain interactions resemble those between ClpS and the ClpA N domain. Glu 184 of Meca and Glu 79 of ClpS have a similarly key role in binding to the N domains of ClpC and ClpA, respectively.

Meca abrogated its interaction with the ClpC N domain (Supplementary Fig. 4a) and removed the ability of Meca to activate ClpC or to form the Meca-ClpCP protease²⁴. The mutation T31A in ClpC also severely compromised its ability to form a hexameric complex with Meca (data not shown). Using *in vitro* assays, we demonstrated that, while the WT Meca-ClpC complex hydrolysed ATP, unfolded the substrate protein GFP-ComK, and in the presence of ClpP led to substrate degradation (Supplementary Fig. 4b-d), the mutations Thr 31 Ala in ClpC and Glu 184 Arg in Meca markedly reduced these activities (Fig. 1d). The mutations Thr 81 Ala in ClpC and Asn 177 Ala in Meca also led to reduced activities.

Unexpectedly, the interactions between Meca121 and the N domain of ClpC resemble those between ClpS and the N domain of ClpA^{25,26} (Fig. 1e). Both Meca and ClpS primarily use a single α -helix to interact with a similar region of the N domains of ClpC and ClpA, respectively. Strikingly, Glu 184 of Meca, which accepts three hydrogen bonds from Thr 31 and Thr 81 of ClpC, occupies the same position as Glu 79 of ClpS, which forms two hydrogen bonds with Glu 28 and Thr 81 of ClpA. The mimicry of interaction is echoed by the finding that ClpS directly interacted with ClpC in *Synechococcus elongatus*²⁷, which, coincidentally, has no obvious Meca orthologue.

Structure of a hexameric Meca-ClpC complex

To understand how Meca facilitates assembly of the ClpC hexamer, we crystallized a stoichiometric complex between Meca121 and a ClpC

variant with D2 deleted (ClpC- Δ D2; residues 1–485, Glu 280 Ala, Δ 247–251 and Δ 281–292). The structure was determined at a resolution of 3.65 Å (Supplementary Table 2 and Supplementary Figs 5 and 6a). Six molecules of Meca121 and six molecules of ClpC- Δ D2 assemble into a basket-shaped complex with a height of 75 Å and a maximum diameter of 150 Å (Fig. 2a). At the wide end of the basket, six molecules of Meca121 interdigitate with six ClpC N domains to form a closed ring that is distal to the ClpP protease. At the narrow end of the basket, six D1 domains form a second ring, through an interface that is similar to that of other AAA+ proteins such as NSF^{28,29} (Fig. 2b and Supplementary Fig. 6b). The six M domains are located on the exterior of the basket, connecting the two closed rings. These M domains, each comprising a pair of coiled coils, emanate from the D1 ring (Fig. 2c) and reach out to interact with Meca121 of the distal ring.

The general features of the hexameric complex of ClpC- Δ D2 are consistent with those of the electron-microscopy-based model of ClpB^{11,12} and are incompatible with those suggested for Hsp104^{14,15}. Specifically, the position of the M domain and the packing interface between adjacent D1 subunits in our X-ray structure are similar to those proposed for ClpB¹². By contrast, the M domains of Hsp104 were placed mainly within the D1 ring, separating two neighbouring D1 domains^{14,15}. This organization results in an unusual packing pattern for D1 and D2 of Hsp104, which differs from the conserved packing interactions for known AAA+ proteins.

The M domain of ClpC contains 58 amino acids (411–468); by contrast, the M domain of ClpB contains 110 residues (Supplementary

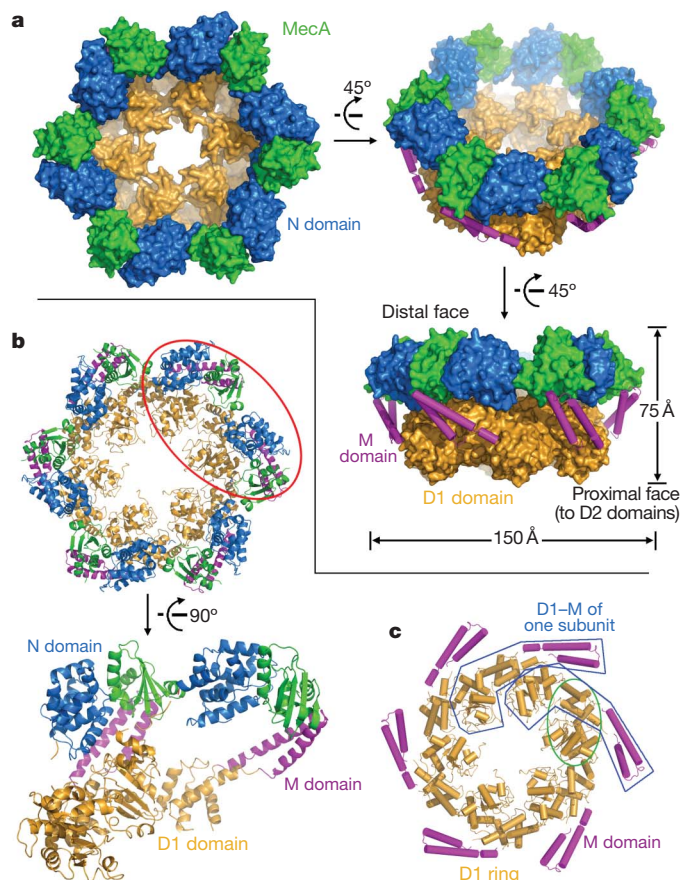


Figure 2 | Structure of the hexameric complex between Meca121 and ClpC- Δ D2. **a**, Surface representation of the hexameric Meca121-ClpC- Δ D2 complex in three related views. **b**, The D1 ring does not directly interact with the distal ring. Their interactions are bridged exclusively by the M domains. **c**, Close-up view of the structure of the D1-M ring. Note the elongated organization of the D1-M domains from one ClpC molecule.

Fig. 7). The M domain of ClpC corresponds to the N-terminal half of the M domain in ClpB, known as motif 1, and lacks the C-terminal half, or motif 2, which directly interacts with the large domain of D1 in the structure of monomeric ClpB¹¹. Superposition of the D1 domains of ClpB and ClpC revealed that their N domains are separated by a distance of approximately 50 Å and that the M domains are at an angle of about 60°.

Interface between MecA and ClpC

A single ClpC-ΔD2 molecule has an extended architecture (Fig. 3a). Formation of the hexameric complex involves four discrete interfaces between MecA and ClpC: two in the distal ring, one in the D1 ring, and one between the M domain and MecA121 (Fig. 3b). In the distal ring, six heterodimers of MecA121–ClpC N domain interact with each other through polar contacts. Glu 198 and Tyr 199 of MecA are within hydrogen-bonding distance of Arg 9, Thr 105 and Glu 106 in the N domain of ClpC (Fig. 3b). The three mutations Glu 198 Ala, Tyr 199 Ala and Lys 201 Asp in MecA weaken its binding to ClpC, abrogate the assembly of the MecA–ClpC protease²⁴ and fail to hydrolyse ATP or unfold substrate (Fig. 3c). The two mis-sense mutations Arg 9 Glu and Glu 106 Arg also compromise the ability of ClpC to form a hexameric complex with MecA (Supplementary Fig. 8) and markedly reduce ATPase activity and substrate unfolding in the presence of MecA (Fig. 3c).

Three charged amino acids, Arg 198, Arg 199 and Asp 396, which are invariant among ClpA, ClpB/Hsp104 and ClpC, map to the interface between two adjacent D1 domains (Fig. 3b). Arg 332, which corresponds to the arginine finger, is also located at this interface. A

double mutation of these amino acids, Arg 198 Ala/Arg 199 Ala, and the individual mutations Arg 332 Ala and Asp 396 Ala compromise formation of the hexameric MecA–ClpC complex (Supplementary Fig. 8) and nearly abrogate ATPase activity and substrate unfolding (Fig. 3d). The invariant nature of these amino acids suggests a conserved packing interface among the Clp/Hsp100 family proteins.

The M domain binds to MecA through a mixture of van der Waals contacts and polar interactions (Fig. 3b). Notably, Phe 436 of ClpC is nestled in a hydrophobic pocket formed by seven MecA residues: Phe 136, Ile 203, Ile 204, Thr 211, Ile 212, His 215 and Phe 216. These interactions are surrounded by a number of polar and charged amino acids, Glu 137, Ser 156 and Tyr 161 of MecA and Asp 428, Gln 432 and Arg 443 of ClpC. The mutation Phe 436 Ser in ClpC abolishes formation of the hexameric MecA–ClpC complex (Supplementary Fig. 8) and abrogates the ATPase and substrate unfolding activities (Fig. 3e). By contrast, the mutation Asp 428 Ala in ClpC has little impact on the assembly of the MecA–ClpC hexamer (Supplementary Fig. 8) but markedly reduces the ATPase activity and substrate unfolding (Fig. 3e). Two double mutations in MecA, Asp 135 Lys/ Glu 13 Lys and Ile 203 Arg/Ile 204 Arg, abrogate ATPase activity, substrate unfolding and ClpP-dependent substrate degradation.

Architecture of the MecA–ClpC molecular machine

After numerous trials, we generated crystals of the hexameric complex between MecA108 and the full-length ClpC (Glu 280 Ala, Glu 618 Ala, Δ247–251, Δ281–292, Δ587–596 and Δ665–685). The structure was determined at a resolution of 6.9 Å by molecular replacement (Supplementary Table 2 and Supplementary Fig. 9). The MecA–ClpC

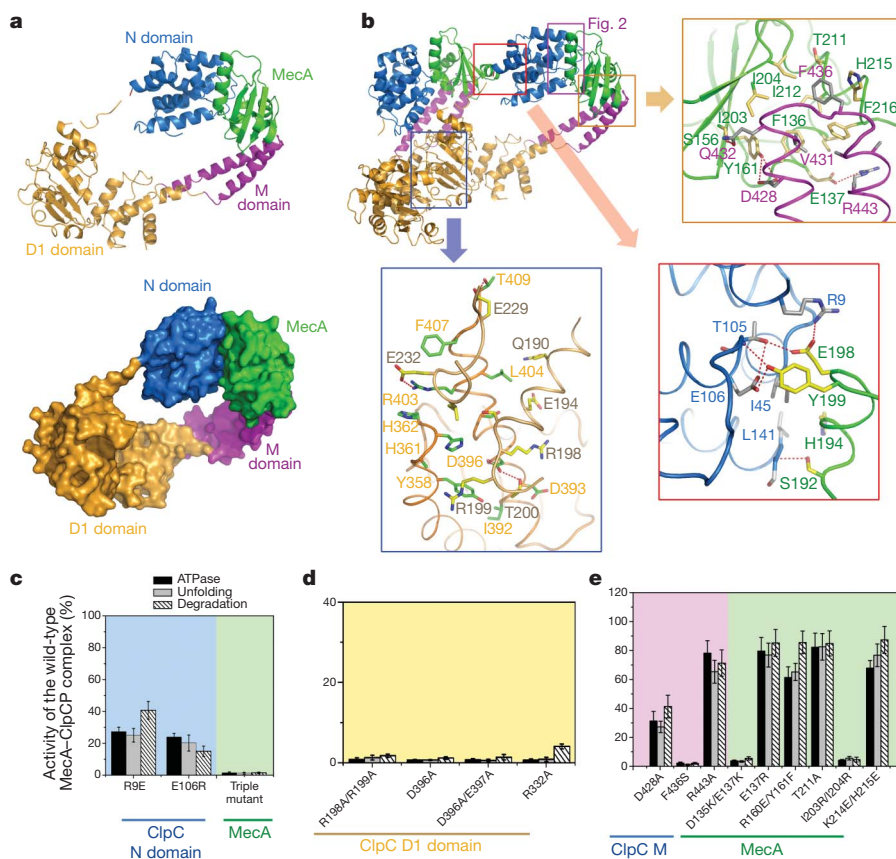


Figure 3 | Recognition interface between MecA121 and ClpC-ΔD2.

a, Structure of a single ClpC-ΔD2 molecule bound to MecA121, shown as a ribbon diagram (top) and in surface view (bottom). **b**, Three of the four interfaces between MecA121 and ClpC-ΔD2, shown in close-up. **c**, Mutational analysis of the interface between MecA121 and the N domain of ClpC. The

three mutations in MecA are Glu 198 Ala, Tyr 199 Ala and Lys 201 Asp.

d, Mutational analysis of the interface between two adjacent D1 domains of ClpC. **e**, Mutational analysis of the interface between MecA121 and the M domain of ClpC. Error bars, s.d. of three sets of experiments.

molecular machine measures 120 Å in height and 150 Å in diameter and comprises three closed rings (Fig. 4a). Each ClpC molecule has an extended appearance, with its N domain, D1 and D2 separated from each other (Fig. 4b). The overall size of the D2 ring, with an exterior diameter of approximately 135 Å, is greater than that of the D1 ring (Fig. 4c). At the resolution limit of 6.9 Å, the six D2 domains, which are responsible for binding to the ClpP protease, seem to be less tightly packed together in comparison with the D1 ring. In addition, the six D2 domains have an asymmetric appearance. Such asymmetry, which is more striking in the hexameric ClpX structure³⁰, is likely to be essential for the function of the D2 ring.

Biochemical analysis

We performed structure-guided mutational analysis on the two nucleotide-binding domains. Mutation of the Walker B residue Glu 280 Ala in ClpC reduces the ATPase activity by 22% (Fig. 4d). By contrast, ATPase reduction by the Walker B mutation Glu 618 Ala in D2 is 85%, suggesting that the D2 domain may contribute to the majority of ATPase activity. Neither of the two Walker B mutations affects assembly of the hexameric Meca–ClpC complex (Supplementary

Fig. 10). Despite retaining the bulk of ATPase activity, ClpC–Glu 280 Ala nearly abrogates the ability to unfold GFP–ComK in the presence of Meca. Together with the fact that ClpC–Glu 280 Ala still recognizes the substrate (data not shown), these observations suggest that the D1 domains probably have a major role in substrate unfolding. Supporting this analysis is the fact that although ClpC–Glu 618 Ala retained only about 15% ATPase activity, it exhibits 50% of the wild-type substrate unfolding activity (Fig. 4d). Nonetheless, ClpC–Glu 618 Ala has only 10% of the wild-type substrate degradation, suggesting that the D2 ring may facilitate translocation of the unfolded substrate protein into the ClpP protease. In contrast to the Walker B mutations, the two Walker A mutations Lys 214 Gln and Lys 551 Gln compromise the assembly of the hexameric Meca–ClpC complex (Supplementary Fig. 10); these mutants have sharply reduced ATPase activities.

Next we examined the impact of deleting pore loops in D1 and D2. The Clp/Hsp100 proteins contain two invariant pore loops, the Lys–Tyr–Arg–Gly loop³¹ (pore loop 1) in D1 and the Gly–Tyr–Val–Gly loop in D2 (Supplementary Fig. 5). D1, but not D2, also contains a second conserved pore loop (pore loop 2), Gly–Ala–Gly–Gly–Ala. Deletion of either pore loop in D1 reduces ATPase activity by 80% and abrogate substrate unfolding and degradation (Fig. 4e). By contrast, deletion of the Gly–Tyr–Val–Gly loop leads only to reduction of ATPase activity, substrate unfolding and degradation. Removal of the ClpP-binding loop in ClpC has little impact on ATPase activity or substrate unfolding but abolishes substrate degradation. All these mutants retain the same ability as wild-type ClpC to form a stable hexameric complex with Meca (data not shown). Except for the ClpC mutant with the ClpP-binding loop deleted, all other ClpC mutants retain interactions with ClpP (data not shown).

To pinpoint the key amino acids, we generated three ClpC variants: Tyr 253 Ala in the Lys–Tyr–Arg–Gly loop, which is thought to contact the substrate polypeptide chain³², Ala 287 Asp/Gly 288 Asp in the Gly–Ala–Gly–Gly–Ala loop and Tyr 593 Ala in the Gly–Tyr–Val–Gly loop. Although ClpC–Tyr 253 Ala and ClpC–Ala 287 Asp/Gly 288 Asp have quite different impacts on ATPase activity, both mutations abolish unfolding and degradation of substrate protein (Fig. 4f). By contrast, ClpC–Tyr 593 Ala reduces, but fails to abrogate, substrate unfolding and degradation. These observations confirm the essential role of the pore loops in the D1 ring. On the basis of the biochemical data and general knowledge of the field, we propose a speculative model for the Meca–ClpC molecular machine (Supplementary Fig. 11).

Discussion

Crystallization of oligomeric AAA+ proteins is hindered by conformational heterogeneity of the individual AAA+ domains, as exemplified by the presence of only one atomic structure of a complete hexameric AAA+ protein containing two nucleotide-binding domains, that of p97^{33–35}. Protein engineering is frequently used to facilitate crystallization by trapping AAA+ proteins in the restrained conformation. In our study, deletion of four internal loops, which nearly abolished ATPase activity and substrate unfolding, proved to be essential for appropriate crystallization. The hexameric complex between Meca121 and ClpC–ΔD2 was crystallized in the absence of exogenous nucleotides. Careful examination of the electron density map failed to reveal any bound nucleotides in ClpC, which was consistent with results of mass spectrometry. The complex between Meca108 and the full-length ClpC was crystallized in the presence of AMP–PNP (5′-adenylyl-imidodiphosphate). Unfortunately, the low resolution, of 6.9 Å, did not allow unambiguous determination of whether AMP–PNP is bound.

Unlike other Clp/Hsp100 proteins, ClpC can form only a functional, hexameric complex in the presence of Meca¹⁷. Our crystal structure reveals the mechanism by which Meca facilitates the hexamerization of ClpC. Meca interacts with ClpC through three interfaces, involving a buried surface area of 2,400 Å² for each Meca molecule. These interactions are expected to strengthen drastically the otherwise transient interactions within the D1 and D2 rings of ClpC.

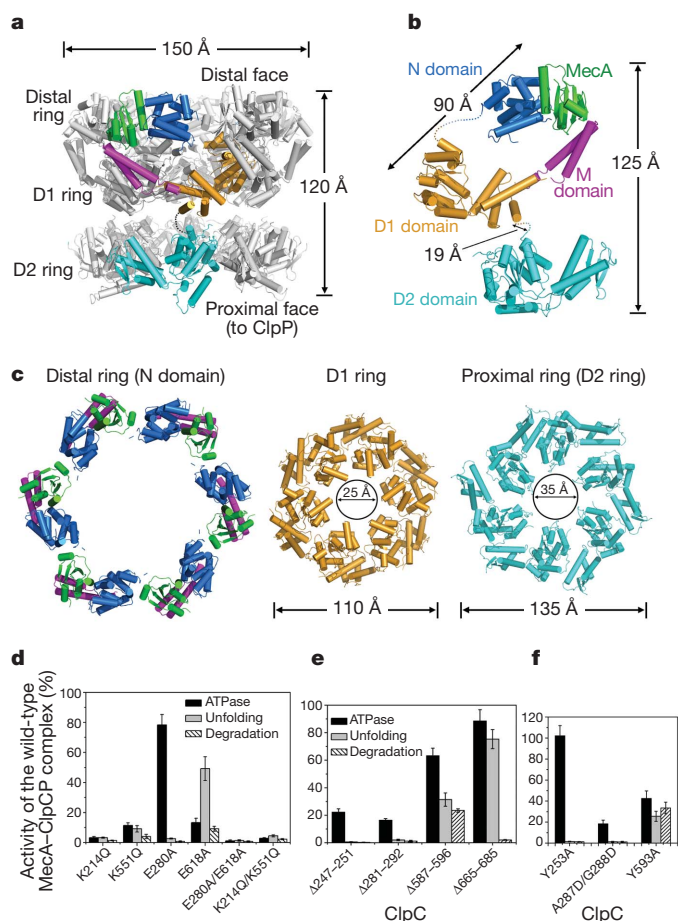


Figure 4 | Architecture of the Meca–ClpC molecular machine. **a**, Overall structure of the hexameric complex between Meca108 and full-length ClpC. One ClpC molecule and the associated Meca are colour-coded; all other molecules are shown in grey. **b**, Structure of a single ClpC molecule bound to Meca108 has an extended organization. **c**, Close-up comparison of the three closed rings in the Meca–ClpC complex. **d**, Functional consequences of mutations that target the Walker A and Walker B motifs. **e**, Functional consequences of pore loop deletions in the D1 and D2 domains of ClpC. The D1 domain contains pore loop 1 (Lys 252–Tyr 253–Arg 254–Gly 255) and pore loop 2 (Gly 286–Ala 287–Gly 288–Gly 289–Ala 290), whereas the D2 domain only has one pore loop (Gly 592–Tyr 593–Val 594–Gly 595). **f**, Functional consequences of mis-sense mutations in the pore loops of the D1 and D2 domains. Error bars, s.d. of three sets of experiments.

METHODS SUMMARY

We expressed and purified ClpC and MecA variants, ClpP and ComK as previously described^{21,24}. To improve crystal quality, we deleted four loops in ClpC, two in D1 (residues 247–251 and 281–292) and two in D2 (residues 587–596 and 665–685), and introduced Walker B mutations Glu 280 Ala and Glu 618 Ala in the Walker B motifs of the D1 and D2 domains, respectively. Assays for ATPase activity, unfolding and degradation were performed as previously described²¹. We grew all crystals by the hanging-drop vapour diffusion method. The data sets were processed using Bruker PROTEUM2³⁶ and XDS³⁷, HKL2000³⁸ and the CCP4 suite³⁹. We determined the heavy-atom positions for crystals of the ClpC N domain alone using SHELXD⁴⁰ and refined them using PHASER⁴¹. All other structures were determined by molecular replacement. All structures were refined using PHENIX⁴².

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 July; accepted 22 December 2010.

Published online 2 March 2011.

- Baker, T. A. & Sauer, R. T. ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends Biochem. Sci.* **31**, 647–653 (2006).
- Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* **78**, 477–513 (2009).
- Goldberg, A. L. Functions of the proteasome: from protein degradation and immune surveillance to cancer therapy. *Biochem. Soc. Trans.* **35**, 12–17 (2007).
- Kirstein, J., Moliere, N., Dougan, D. A. & Turgay, K. Adapting the machine: adaptor proteins for Hsp100/Clp and AAA+ proteases. *Nature Rev. Microbiol.* **7**, 589–599 (2009).
- Inobe, T. & Matouschek, A. Protein targeting to ATP-dependent proteases. *Curr. Opin. Struct. Biol.* **18**, 43–51 (2008).
- Groll, M., Bochtler, M., Brandstetter, H., Clausen, T. & Huber, R. Molecular machines for protein degradation. *ChemBioChem* **6**, 222–256 (2005).
- Barends, T. R., Werbeck, N. D. & Reinsteiner, J. Disaggregases in 4 dimensions. *Curr. Opin. Struct. Biol.* **20**, 46–53 (2010).
- Doyle, S. M. & Wickner, S. Hsp104 and ClpB: protein disaggregating machines. *Trends Biochem. Sci.* **34**, 40–48 (2009).
- Haslberger, T., Bukau, B. & Mogk, A. Towards a unifying mechanism for ClpB/Hsp104-mediated protein disaggregation and prion propagation. *Biochem. Cell Biol.* **88**, 63–75 (2010).
- Lee, S., Sowa, M. E., Choi, J. M. & Tsai, F. T. The ClpB/Hsp104 molecular chaperone—a protein disaggregating machine. *J. Struct. Biol.* **146**, 99–105 (2004).
- Lee, S. *et al.* The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. *Cell* **115**, 229–240 (2003).
- Lee, S., Choi, J. M. & Tsai, F. T. Visualizing the ATPase cycle in a protein disaggregating machine: structural basis for substrate binding by ClpB. *Mol. Cell* **25**, 261–271 (2007).
- Lee, S., Sielaff, B., Lee, J. & Tsai, F. T. CryoEM structure of Hsp104 and its mechanistic implication for protein disaggregation. *Proc. Natl Acad. Sci. USA* **107**, 8135–8140 (2010).
- Wendler, P. *et al.* Atypical AAA+ subunit packing creates an expanded cavity for disaggregation by the protein-remodeling factor Hsp104. *Cell* **131**, 1366–1377 (2007).
- Wendler, P. *et al.* Motor mechanism for protein threading through Hsp104. *Mol. Cell* **34**, 81–92 (2009).
- Wendler, P. & Saibil, H. R. Cryo electron microscopy structures of Hsp100 proteins: crowbars in or out? *Biochem. Cell Biol.* **88**, 89–96 (2010).
- Kirstein, J. *et al.* Adaptor protein controlled oligomerization activates the AAA+ protein ClpC. *EMBO J.* **25**, 1481–1491 (2006).
- Persuh, M., Turgay, K., Mandic-Mulec, I. & Dubnau, D. The N- and C-terminal domains of MecA recognize different partners in the competence molecular switch. *Mol. Microbiol.* **33**, 886–894 (1999).
- Turgay, K., Hamoen, L. W., Venema, G. & Dubnau, D. Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of *Bacillus subtilis*. *Genes Dev.* **11**, 119–128 (1997).
- Turgay, K., Hahn, J., Burghoorn, J. & Dubnau, D. Competence in *Bacillus subtilis* is controlled by regulated proteolysis of a transcription factor. *EMBO J.* **17**, 6730–6738 (1998).
- Mei, Z. *et al.* Molecular determinants of MecA as a degradation tag for the ClpCP protease. *J. Biol. Chem.* **284**, 34366–34375 (2009).
- Guo, F., Maurizi, M. R., Esser, L. & Xia, D. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. *J. Biol. Chem.* **277**, 46743–46752 (2002).
- Li, J. & Sha, B. Crystal structure of the *E. coli* Hsp100 ClpB N-terminal domain. *Structure* **11**, 323–328 (2003).
- Wang, F. *et al.* Crystal structure of the MecA degradation tag. *J. Biol. Chem.* **284**, 34376–34381 (2009).
- Guo, F., Esser, L., Singh, S. K., Maurizi, M. R. & Xia, D. Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. *J. Biol. Chem.* **277**, 46753–46762 (2002).
- Zeth, K. *et al.* Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nature Struct. Biol.* **9**, 906–911 (2002).
- Andersson, F. I. *et al.* Cyanobacterial ClpC/HSP100 protein displays intrinsic chaperone activity. *J. Biol. Chem.* **281**, 5468–5475 (2006).
- Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weiss, W. I. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94**, 525–536 (1998).
- Yu, R. C., Hanson, P. I., Jahn, R. & Brunger, A. T. Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nature Struct. Biol.* **5**, 803–811 (1998).
- Glynn, S. E., Martin, A., Nager, A. R., Baker, T. A. & Sauer, R. T. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell* **139**, 744–756 (2009).
- Hinnerwisch, J., Fenton, W. A., Furtak, K. J., Farr, G. W. & Horwich, A. L. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell* **121**, 1029–1041 (2005).
- Schlieker, C. *et al.* Substrate recognition by the AAA+ chaperone ClpB. *Nature Struct. Mol. Biol.* **11**, 607–615 (2004).
- DeLaBarre, B. & Brunger, A. T. Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains. *Nature Struct. Biol.* **10**, 856–863 (2003).
- DeLaBarre, B. & Brunger, A. T. Nucleotide dependent motion and mechanism of action of p97/VCP. *J. Mol. Biol.* **347**, 437–452 (2005).
- Davies, J. M., Brunger, A. T. & Weiss, W. I. Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. *Structure* **16**, 715–726 (2008).
- Bruker. S.A.I.N.T., SADABS, XPREP and SHELXTL/NT Software Reference Manual (Bruker AXS, 2003).
- Kabsch, W. Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallogr.* **21**, 916–924 (1988).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772–1779 (2002).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
- Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
- DeLano, W. L. PyMOL Molecular Viewer (<http://www.pymol.org>) (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. He and S. Huang at Shanghai Synchrotron Radiation Source and staff at the SPring-8 beamline BL41XU for help. This work was supported by Project 30888001 of the National Natural Science Foundation of China (Y.S.) and Postdoctoral Fellowship #023201058 from the China Postdoctoral Science Foundation (Z.M.).

Author Contributions F.W., Z.M., and Y.S. designed all experiments. F.W., Z.M., Y.Q., C.Y., Q.H., and J.W. performed the experiments. All authors analysed the data. F.W., Z.M., J.W. and Y.S. contributed to manuscript preparation. Y.S. wrote the manuscript.

Author Information The atomic coordinates and structure factors of the ClpC N domain, the MecA121–ClpC N-domain complex, the hexameric complex of MecA121–ClpC-ΔD2 and the hexameric complex of MecA108–ClpC have been deposited in the Protein Data Bank with the accession codes 2Y1Q, 2Y1R, 3PXG and 3PXL, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.W. (jwwang@tsinghua.edu.cn) or Y.S. (shi-lab@tsinghua.edu.cn).

METHODS

Protein preparation. All clones were generated using a standard PCR-based cloning strategy and the identities of individual clones were verified through double-strand plasmid sequencing. ClpC and MecA variants, and ClpP and ComK used for biochemical assays, were expressed and purified as described previously^{21,24}. In addition, GFP–ComK for unfolding assays was overexpressed in *Escherichia coli* strain BL21 (DE3) using the PLM303 vector (a derivative of pET-27a) and purified as described for ComK^{21,24}.

For crystallization, the ClpC N domain and the stoichiometric complex between MecA121 and the ClpC N domain were obtained using a Superdex-200 column (10/30, GE Healthcare). To improve the X-ray diffraction limit of the crystals for the complexes between MecA121 and ClpC-ΔD2 or full-length ClpC, we deleted two or four internal loop sequences, respectively, in the D1 domain (residues 247–251 and 281–292) or D1/D2 domains (residues 247–251, 281–292, 587–596 and 665–685). The selection of these loops was based on sequence alignment among ClpA, ClpB and ClpC; the ClpC sequences that correspond to the surface loops that are disordered in the ClpA/ClpB monomer structures were presumed to affect crystal packing negatively. Indeed, such practice resulted in marked improvement of the X-ray diffraction limit. To prevent potential problems associated with ATP hydrolysis, we introduced mis-sense mutations (Glu280 Ala, Glu618 Ala) in the Walker B motif in the D1 and D2 domains. MecA121 and ClpC-ΔD2 or full-length ClpC were co-expressed in *E. coli* BL21(DE3) using the vectors pGEX-6P-1 (GE Healthcare) and pBB75 (Novagen), respectively. The complexes were purified over GS4B resin (GE Healthcare) and cleaved by PreScission Protease (GE Healthcare), followed by anion-exchange chromatography (Source-15Q, GE Healthcare) and gel filtration (Superdex-200). The concentrations of proteins were determined by spectroscopic measurement at 280 nm. All recombinant proteins were characterized by gel filtration.

ATPase assay. The ATPase activity of ClpC was determined in the presence of MecA using a coupled spectrophotometric assay⁴⁴. All reactions were carried out at 37 °C in 1.8 ml of reaction mixture containing 100 mM HEPES, pH 7.5, 100 mM KCl, 20 mM ATP, 20 mM MgCl₂, 1 mM NADH, 2 mM phosphoenolpyruvate, 50 units ml^{−1} lactate dehydrogenase and 50 units ml^{−1} pyruvate kinase. In these reactions, ClpC variants and MecA variants were fixed at 2 μM and 3 μM, respectively. The reaction was monitored continuously by following the formation of NAD⁺ at 340 nm on a PerkinElmer Life Sciences Lambda 45 spectrophotometer equipped with a magnetic stirrer in the cuvette holder. The concentrations of ADP generated were determined using a molar extinction coefficient of 6,220 cm^{−1} M^{−1} for NADH at 340 nm. The concentrations of all proteins were determined spectrophotometrically on the basis of their molar extinction coefficients calculated from their amino-acid sequences.

Unfolding assay. The unfolding assay was modified from a published protocol⁴⁵. Fluorescence measurement of GFP, with an excitation wavelength of 400 nm and an emission wavelength of 510 nm, was performed on a Hitachi fluorescence spectrometer F4600. The unfolding reactions were carried out at 37 °C in 100 mM HEPES, pH 7.5, 100 mM KCl, 20 mM ATP, 20 mM MgCl₂, 2 mM phosphoenolpyruvate and 50 units ml^{−1} pyruvate kinase. In these reactions, ClpC variants and MecA variants were added at 2 μM and 3 μM, respectively. GFP–ComK was used at 6 μM. We serendipitously discovered that, once unfolded, GFP–ComK precipitated out of solution and was unable to refold spontaneously. Consequently, unfolding assays performed in the presence and absence of the GroEL-trap mutant⁴⁵ gave rise to identical results for the first 100 s; these were used to calculate the unfolding rate. Precipitation of the unfolded GFP–ComK protein did not affect measurement of the fluorescence signal. These observations allowed us to perform the unfolding assays in the absence of the GroEL-trap mutant.

ClpCP degradation assay. *In vitro* degradation assays were carried out as described previously^{19,20}. ClpC and ClpP were added at a final concentration of 4 μM. All MecA variants and ComK or GFP–ComK were added at 6 μM final concentration. The bands on SDS–polyacrylamide gel electrophoresis (SDS–PAGE) were quantified with Quantity One software (Bio-Rad) to determine the substrate degradation rate. All protein concentrations reported in this manuscript refer to those of the monomers, regardless of the oligomerization states of the proteins.

Interaction assay by gel filtration. Size exclusion chromatography, using a Superdex-200 column (10/30, GE Healthcare), was used to examine protein interactions. In all cases, proteins were incubated at 4 °C for at least 45 min to allow equilibrium to be reached. The flow rate was 0.5 ml min^{−1}, and the buffer contained 10 mM Tris, pH 8.0, 150 mM NaCl, 2 mM dithiothreitol and 2 mM ATP (unless otherwise indicated). All fractions were collected at 0.5 ml each. Aliquots of relevant fractions were mixed with SDS sample buffer and subjected to SDS–PAGE. The proteins were visualized by Coomassie blue staining. The column was calibrated with molecular mass standards.

Crystallization. Crystals were grown by the hanging-drop vapour diffusion method by mixing an equal volume of sample (10–15 mg ml^{−1}) in 10 mM Tris–HCl, pH 8.0, and 150 mM NaCl with reservoir solution at 18 °C. Crystals of the ClpC N domain were grown from the reservoir buffer containing 2.0 M (NH₄)₂SO₄ and 100 mM sodium acetate, pH 4.6. Crystals of MecA121 bound to ClpC N domain were grown from the reservoir buffer that contained 20% (w/v) PEG 3350, 300 mM ammonium tartrate and 6 mM *N*-nonyl-β-D-maltopyranoside.

We prepared a stoichiometric complex between MecA121 and the wild-type ClpC protein with the D2 domain deleted (residues 1–485, referred to as ClpC-ΔD2). Introduction of a walker B mutation Glu 280 Ala allowed crystallization of this complex under several conditions. However, none of these crystals diffracted X-rays beyond a resolution of 5 Å. Reasoning that the poor diffraction might be caused by flexible surface loops, we generated a large number of ClpC variants, each containing deletion of one or two surface loops. Serendipitously, removal of residues 247–251 and 281–292 markedly improved the crystal diffraction limit. MecA121 in complex with ClpC-ΔD2 (Glu 280 Ala, Δ247–251, Δ281–292) was crystallized in the reservoir solution containing 12% PEG 3350, 450 mM sodium malonate and 100 mM MES, pH 7.0.

After numerous trials, we generated crystals of the hexameric complex between MecA121 and full-length ClpC with Walker B mutations Glu 280 Ala/Glu 618 Ala, which unfortunately suffered from the recurring problem of poor diffraction. Taking advantage of the lessons learned from our prior effort, we were able to improve the X-ray diffraction limit to approximately 8 Å by removing four internal loops, residues 247–251, 281–292, 587–596 and 665–685. Deletion of residues 665–685 resulted in the removal of the ClpP-binding loop⁴⁶. A longer version of MecA (residues 108–218, hereafter referred to as MecA108) further improved the resolution limit to about 6.9 Å. To obtain the crystal of MecA108 bound to full-length ClpC (Glu 280 Ala, Glu 618 Ala, Δ247–251, Δ281–292, Δ587–596 and Δ665–685), 5 mM AMP–PNP was added to the protein solution and incubated before crystallization. Crystals were grown in the drops against the reservoir buffer containing 10% PEG 3350, 400 mM magnesium formate and 100 mM MES, pH 6.9.

X-ray data collection. The native data and the iodine-soaked single-wavelength anomalous dispersion (SAD) data of ClpC N domain and the native data of the binary complex between MecA121 and ClpC N domain were collected on our home X-ray diffractometers. These data sets were processed using Bruker PROTEUM2³⁶ and XDS³⁷. The native data of the complex between MecA121 and ClpC-ΔD2 (Glu 280 Ala, Δ247–251, Δ281–292) and between MecA108 and full-length ClpC (Glu 280 Ala, Glu 618 Ala, Δ247–251, Δ281–292, Δ587–596 and Δ665–685) were collected at the SPring-8 beamline BL41XU and the Shanghai Synchrotron Radiation Source. Because the long dimension of the unit cell is much larger than the other two for these two crystals, a bent cryo-loop was used to orient the long axis of the crystals to be more or less parallel to the goniostat spindle axis, to avoid overlap of the diffraction spots. The synchrotron data sets were integrated and scaled using the HKL2000 package³⁸. Further data processing was carried out using programs from the CCP4 suite³⁹. Data collection statistics are summarized in Supplementary Tables 1 and 2.

Structural determination. The iodine positions in the NaI-derived crystal of ClpC N domain were determined using the program SHELXD⁴⁰. The identified iodine sites were refined and the initial phases were generated in the program PHASER⁴¹ with the SAD experimental phasing module. Solvent flattening and histogram matching were performed using DM⁴⁷. The initial model was traced automatically using the program BUCCANEER⁴⁸, and was manually rebuilt in COOT⁴⁹. The final structure was refined with PHENIX⁴².

The structure of the binary complex between MecA121 and ClpC N domain was determined with molecular replacement using the models of ClpC N domain from above and MecA121 (PDB ID, 3JTP²⁴). The structure of the complex between MecA121 and ClpC-ΔD2 (Glu 280 Ala, Δ247–251, Δ281–292) was solved with molecular replacement using the atomic coordinates of the complex between MecA121 and ClpC N domain and the ClpB D1 model (PDB ID, 1QVR¹¹). The electron density that corresponds to residues 143–154 was weak and did not allow unambiguous assignment of side chains. All these structures were refined using PHENIX⁴² with restraints, including stereochemistry, non-crystallographic symmetry and main-chain hydrogen bonds.

The Matthews coefficient of the complex crystal between MecA108 and full-length ClpC (Glu 280 Ala, Glu 618 Ala, Δ247–251, Δ281–292, Δ587–596 and Δ665–685) indicates that only half of the hexamer is present in one asymmetric unit. Half of the hexameric complex between MecA121 and ClpC-ΔD2 (Glu 280 Ala, Δ247–251, Δ281–292) was used as the search model to solve this structure with molecular replacement. One obvious solution was obtained with the space group *P*6₅22. Next, the atomic coordinates of the ClpC D2 domain (Qi *et al.*, work in progress) was also searched with the above MecA121 and ClpC-ΔD2 coordinates fixed. Three D2 molecules were found despite the relatively low

resolution of the data, 6.9 Å. The final structure was refined with PHENIX⁴², only using the rigid-body refinement of individual domains and group ADP parameter for the whole domain.

44. Cook, P. F., Neville, M. E. Jr, Vrana, K. E., Hartl, F. T. & Roskoski, R. Jr. Adenosine cyclic 3',5'-monophosphate dependent protein kinase: kinetic mechanism for the bovine skeletal muscle catalytic subunit. *Biochemistry* **21**, 5794–5799 (1982).
45. Weber-Ban, E. U., Reid, B. G., Miranker, A. D. & Horwich, A. L. Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. *Nature* **401**, 90–93 (1999).
46. Kim, Y. I. *et al.* Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nature Struct. Biol.* **8**, 230–233 (2001).
47. Cowtan, K. DM: an automated procedure for phase improvement by density modification. *Joint CCP4 ESF-EACBM Newslett. Protein Crystallogr.* **31**, 34–38 (1994).
48. Cowtan, K. The Buccaneer software for automated model building. *Acta Crystallogr. D* **62**, 1002–1011 (2006).
49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

Crystal structure of a potassium ion transporter, TrkH

Yu Cao^{1*}, Xiangshu Jin^{2*}, Hua Huang^{1*}, Mehabaw Getahun Derebe³, Elena J. Levin¹, Venkataraman Kabaleeswaran¹, Yaping Pan¹, Marco Punta^{4,5}, James Love⁴, Jun Weng¹, Matthias Quick^{6,7}, Sheng Ye³, Brian Kloss⁴, Renato Bruni⁴, Erik Martinez-Hackert⁸, Wayne A. Hendrickson⁸, Burkhard Rost^{4,5}, Jonathan A. Javitch^{6,7,9}, Kanagalaghatta R. Rajashankar¹⁰, Youxing Jiang³ & Ming Zhou¹

The TrkH/TrkG/KtrB proteins mediate K⁺ uptake in bacteria and probably evolved from simple K⁺ channels by multiple gene duplications or fusions. Here we present the crystal structure of a TrkH from *Vibrio parahaemolyticus*. TrkH is a homodimer, and each protomer contains an ion permeation pathway. A selectivity filter, similar in architecture to those of K⁺ channels but significantly shorter, is lined by backbone and side-chain oxygen atoms. Functional studies showed that TrkH is selective for permeation of K⁺ and Rb⁺ over smaller ions such as Na⁺ or Li⁺. Immediately intracellular to the selectivity filter are an intramembrane loop and an arginine residue, both highly conserved, which constrict the permeation pathway. Substituting the arginine with an alanine significantly increases the rate of K⁺ flux. These results reveal the molecular basis of K⁺ selectivity and suggest a novel gating mechanism for this large and important family of membrane transport proteins.

K⁺ is highly concentrated in all living cells and plays diverse physiological roles such as setting membrane potential¹, regulating turgor pressure² and maintaining intracellular pH^{3–5}. Because K⁺ is virtually impermeable to the cell membrane, specialized K⁺ transporters have evolved to mediate its uptake. In animal cells, K⁺ uptake is mainly achieved by the Na⁺/K⁺-ATPase⁶, whereas in non-animal cells, the task is shared by at least two different systems, one of which, the superfamily of K⁺ transporters (SKT proteins)⁷, is the focus of this research.

An SKT protein has four tandem domains with low homology to each other, each resembling a single protomer from a simple K⁺-channel with a predicted M1–P–M2 transmembrane topology^{8–10}. M1 and M2 are transmembrane helices that are connected by P, the re-entrant pore loop, which is composed of a half-membrane-spanning helix followed by an extended loop¹¹. In K⁺ channels, the half-membrane-spanning helix is called the pore helix, and the extended loop harbours the highly conserved signature sequence TVGYG, which forms the selectivity filter responsible for the coordination of K⁺. In SKT proteins, the equivalent sequences in the P-loops are less strictly conserved. It has been proposed that SKT proteins have a structure that resembles K⁺ channels^{9,10,12}. Results from functional studies have been consistent with this view: mutating a conserved glycine residue in the P-loops changes ion selectivity of SKT proteins in bacteria^{7,13} and plants¹⁴; whereas epitope tagging of a yeast SKT protein showed that its transmembrane topology is consistent with four M1–P–M2 repeats¹⁵.

Studies have shown that the selectivity of K⁺ channels is highly sensitive to changes to the structure of the selectivity filter. Introducing point mutations to the signature sequence of K⁺ channels compromises K⁺

selectivity¹⁶; and the NaK channel, which has a slightly modified signature sequence of TVGDG resulting in an altered selectivity filter conformation, is essentially non-selective between K⁺ and Na⁺ (ref. 17). How, then, can the selectivity of the SKT proteins be maintained, given that their signature sequences are so highly degraded (Supplementary Fig. 1)? Furthermore, if SKT proteins have the architecture of a K⁺ channel pore, how is transport of K⁺ controlled? To address these questions we targeted the bacterial TrkG/TrkH/KtrB proteins, the largest sub-family of SKT proteins, for structural and functional studies. The importance of these proteins in bacteria has been shown in *Escherichia coli*, which does not grow in less than 5 mM K⁺ when its *trkG* and *trkH* genes are deleted¹⁸, and in *Francisella tularensis*, which loses its ability to cause tularaemia when lacking TrkH¹⁹. We crystallized and solved the structure of a TrkH from *V. parahaemolyticus* (hereafter VpTrkH).

Function of VpTrkH

VpTrkH was expressed and purified to homogeneity. The purified VpTrkH eluted as a single symmetrical peak on a size-exclusion column (Supplementary Fig. 2a), and by combining the size-exclusion chromatography with light scattering and refractive index measurements, the molecular mass of the purified protein was estimated to be approximately 100 kDa. Because the molecular mass of each VpTrkH protomer is approximately 54 kDa, the detergent-solubilized VpTrkH is a homodimer. The homodimeric quaternary assembly was further verified by running the protein on an SDS–polyacrylamide gel electrophoresis (SDS–PAGE) gel after incubation with covalent cross-linkers, which produced a band roughly twice the molecular mass (Supplementary Fig. 2b). Dimeric assembly has also been reported

¹Department of Physiology & Cellular Biophysics, College of Physicians and Surgeons, Columbia University, 630 West 168th Street, New York, New York 10032, USA. ²Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University, 1130 St Nicholas Avenue, Room 815, New York, New York 10032, USA.

³Department of Physiology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. ⁴New York Consortium on Membrane Protein Structure, New York Structural Biology Center, 89 Convent Avenue, New York, New York 10027, USA. ⁵Department of Computer Science and Institute for Advanced Study, Technical University of Munich, D-85748 Munich, Germany. ⁶Center for Molecular Recognition and Department of Psychiatry, Columbia University, 630 West 168th Street, New York, New York 10032, USA. ⁷New York State Psychiatric Institute, Division of Molecular Therapeutics, 1051 Riverside Drive, New York, New York 10032, USA. ⁸Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University, 630 West 168th Street, New York, New York 10032, USA. ⁹Department of Pharmacology, Columbia University, 630 West 168th Street, New York, New York 10032, USA. ¹⁰Department of Chemistry and Chemical Biology, Cornell University, NE-CAT, Advanced Photon Source, Argonne, Illinois 60439, USA.

*These authors contributed equally to this work.

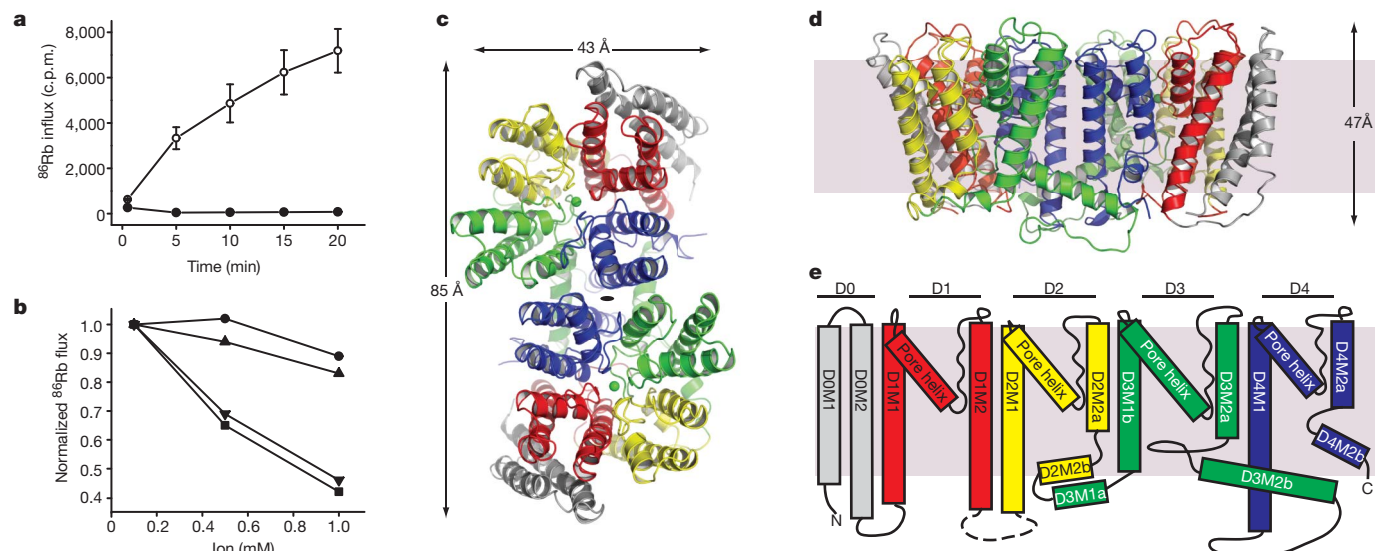


Figure 1 | Function and structure of VpTrkH. **a**, Time-dependent ^{86}Rb influx into liposomes reconstituted with VpTrkH (open circle) or empty vesicles (filled circle). **b**, ^{86}Rb influx at 20 min in the presence of three concentrations of K^+ (square), Rb^+ (inverted triangle), Na^+ (circle) and Li^+ (triangle). **c**, The VpTrkH dimer coloured by domain and viewed from the extracellular side. The twofold symmetry axis is marked as a black oval. The green spheres are K^+ atoms. **d**, VpTrkH viewed from within the membrane

for the KtrB protein from *Bacillus subtilis*^{20,21}, a protein closely related to VpTrkH.

Because VpTrkH had not been functionally characterized, we proceeded to measure its function in two experiments. First, the *VpTrkH* gene was introduced into an *E. coli* strain, LB650 (ref. 18), that lacks both the *trkG* and *trkH* genes. LB650 cells have a slow growth phenotype in media with less than 5 mM K^+ , and expression of VpTrkH rescued the growth of these cells (Supplementary Fig. 2c), suggesting that VpTrkH mediates K^+ uptake. Second, purified and detergent-solubilized VpTrkH protein was reconstituted into liposomes, and K^+ permeation was measured indirectly by monitoring uptake of radioactive $^{86}\text{Rb}^+$ (refs 17, 22). In this experiment, the proteoliposomes contain a high internal concentration of K^+ and are diluted into an external solution with a low concentration of K^+ and a trace amount of radioactive $^{86}\text{Rb}^+$. Efflux of K^+ down its chemical gradient via VpTrkH creates an electrical potential across the bilayer that drives uptake of $^{86}\text{Rb}^+$, an ion that is known to permeate K^+ channels. Uptake of $^{86}\text{Rb}^+$ was observed only in vesicles reconstituted with VpTrkH (Fig. 1a), indicating that VpTrkH is permeable to both K^+ and Rb^+ . Further experiments showed that $^{86}\text{Rb}^+$ uptake is inhibited by external K^+ or Rb^+ but relatively unaffected by the presence of Na^+ or Li^+ in the external solution (Fig. 1b). These results indicate that, despite the weak conservation of its signature sequence, VpTrkH exhibits selectivity for K^+ and Rb^+ over Na^+ and Li^+ . In addition, they also suggest that VpTrkH is capable of mediating facilitated diffusion of K^+ driven by an electrochemical gradient.

Overall structure

Crystals of both native and selenomethionine-substituted VpTrkH were grown under oil by the microbatch method, and the best native protein crystals diffracted to 3.5 Å and belonged to the space group $P2_12_1$ (Supplementary Table 1). Initial phases were obtained by single-wavelength anomalous diffraction²³ using a selenomethionine substituted VpTrkH crystal. The building and refinement of the structural model were facilitated by the presence of a twofold non-crystallographic symmetry (NCS) axis and the positions of the selenium atoms. The final refined model contains residues 1–157, 174–484 and one K^+ per subunit. The region encompassing residues

with the extracellular side on top. The dimer is rotated by 90° about the *x* and *y* axes relative to **c**. Grey rectangle representing the membrane is shown with a thickness of 30 Å. **e**, VpTrkH topology shown with the extracellular side on top. The five domains are coloured according to the same scheme as in the previous panels. The grey rectangle indicates the thickness of the cell membrane, and the unresolved loop is shown as a dashed line.

158–173, which is a loop between two transmembrane helices, is disordered.

Each asymmetric unit contains a dimer of VpTrkH protomers related by a twofold symmetry axis perpendicular to the plane of the membrane. Their amino (N) and carboxy (C) termini both probably reside on the cytoplasmic side as inferred from the experimentally determined topology of a highly homologous TrkH protein from *E. coli*²⁴. Two views of the dimer are shown in Fig. 1c, d. Viewed along the twofold axis from the extracellular side, the dimer has a parallelogram shape with sides of approximately 85 and 43 Å. Along the twofold axis, VpTrkH is approximately 47 Å thick. Stereo views of the VpTrkH dimer in three orientations are shown in Supplementary Fig. 3a–c.

Each VpTrkH protomer is composed of five domains, defined sequentially from the N terminus to the C terminus as D0 to D4 (Fig. 1e and Supplementary Figs 1 and 3b, d). D0, which is found only in the TrkH/TrkG subfamily of SKT proteins, has two transmembrane segments. D1 to D4 each have a K^+ -channel-like M1–P–M2 topology, although the M2 helices of D2 to D4 are composed of two shorter ones (Fig. 2a). The secondary structure of each P-loop also resembles that of a K^+ channel, composed of a half membrane-spanning helix, the pore helix, followed by an extended loop that forms the selectivity filter. D1 to D4 assemble around a pseudo-fourfold symmetry axis to form an ion permeation pathway; when observed from the extracellular side along the pseudo-fourfold axis, they are arranged in an anticlockwise direction (Supplementary Fig. 3b). An extensive dimer interface is composed of helices from D3 and D4, with a buried surface of 2,225 Å² per protomer. There is a hydrophobic cavity in the middle of the interface, but it is sealed off from the aqueous medium by two layers of hydrophobic residues on both sides of the membrane (Supplementary Fig. 4). The ion permeation pathway, which is contained within each protomer, has an hourglass shape and two salient features: a selectivity filter and an intramembrane loop (Fig. 2b).

Selectivity filter

The selectivity filter is surrounded by the four pore helices, which are arranged with the negative ends of their helix-dipole moments pointing to the middle of the membrane (Supplementary Fig. 5). A similar arrangement of pore helices is present in K^+ channels, and its role in

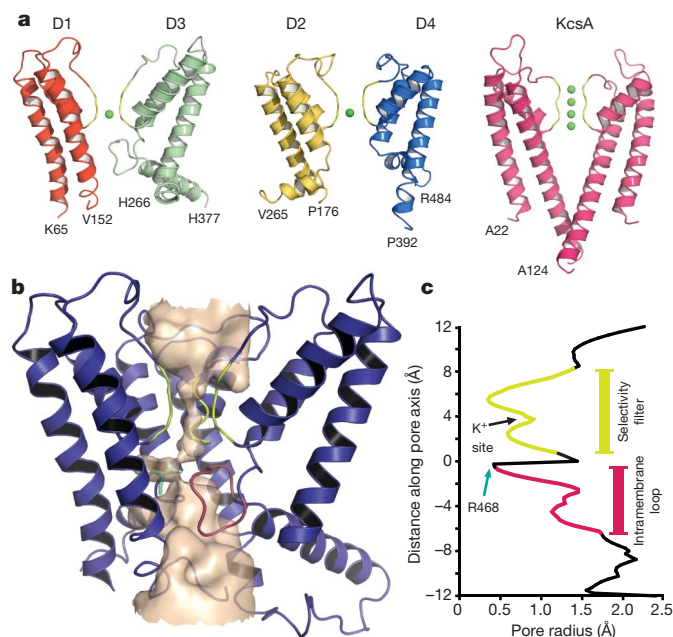


Figure 2 | The VpTrkH pore. **a**, Views of a VpTrkH protomer showing only the D1 and D3 domains (left) or the D2 and D4 domains (middle). Two domains of KcsA are shown on the right for comparison. K^+ atoms are shown as green spheres, and the N- and C-terminal residues are labelled. **b**, Surface representation of the pore of a TrkH subunit obtained with the program Hollow⁴³ using a 1.4-Å probe radius for the vestibules and a 0.75-Å probe radius for the constricted region. The protein is shown with domain 2 removed for clarity and the selectivity filter (yellow), the intramembrane loop (magenta) and residue R468 (teal) highlighted. **c**, Radius of the pore calculated with the program HOLE.

minimizing the free energy of a permeating cation has been discussed for the KcsA K^+ channel^{11,25}.

The four selectivity filter signature sequences are located on the P-loops (Fig. 3a), and these elements come together to form the selectivity filter (omit map in Fig. 3b and Supplementary Fig. 6a, $2F_o - F_c$ map in Supplementary Fig. 6b). Compared with those of K^+ channels, the selectivity filter has a wider opening on the extracellular side and a

much shorter constricted region where K^+ is coordinated (Fig. 3c, d and Supplementary Figs 6c and 7). In the constricted region, main-chain and side-chain oxygen atoms form stacks of oxygen rings that could coordinate and stabilize dehydrated K^+ , a feature that is preserved from the K^+ channels. In an $F_o - F_c$ map calculated with K^+ omitted, two peaks of positive electron density appear in the filter (Fig. 3d and Supplementary Fig. 7a), which we call the upper and lower site and interpret as potential positions where K^+ binds. Owing to the modest resolution of the diffraction data, we cannot unambiguously determine the contribution of K^+ to these densities, as opposed to partial contribution from water molecules or calcium ions that were included in the crystallization solution. Because Rb^+ and Ba^{2+} are known to occupy K^+ positions in the selectivity filter of the KcsA K^+ channel^{11,26}, we took advantage of this knowledge and grew crystals in the presence of Rb^+ or Ba^{2+} (Supplementary Table 1). Difference electron densities corresponding to Fourier coefficients $F_{Rb} - F_K$ or $F_{Ba} - F_K$ are shown in Fig. 3e. In both cases, a strong positive electron density peak is present in the upper site, consistent with substitution of a K^+ with a more electron dense Rb^+ or Ba^{2+} . The upper site lines up with site 3 (S3) in the KcsA K^+ channel, and is constructed entirely by backbone carbonyl oxygen atoms (Fig. 3c and Supplementary Fig. 6c). In the KcsA K^+ channels, Rb^+ does not occupy every K^+ binding site in the selectivity filter, and Rb^+ permeates with a much slower rate than K^+ (refs 27, 28). A slower rate of Rb^+ uptake was also observed by TrkH in *E. coli*²⁹. Therefore, we postulate that the lower peak in the K^+ difference map, which aligns with S4 in KcsA, could potentially be a K^+ binding site (Fig. 3d and Supplementary Fig. 7), although there was no density at this location in either the Rb^+ or Ba^{2+} difference maps.

Although only one binding site has been confirmed by heavy atoms and modelled into the structure, the limited resolution of the data probably reduced our ability to observe more. Comparison with the KcsA structure suggests that the selectivity filter could potentially accommodate a maximum of three K^+ binding sites without requiring a substantial structural change. In addition to the confirmed site and the possible site corresponding to S4 in KcsA, the backbone carbonyls from the highly conserved glycine residues form a ring of oxygen atoms above the confirmed ion-binding site, and could potentially form another K^+ binding site that would line up with S2 in KcsA (Fig. 3d and Supplementary Fig. 7). The S1 site is lost because of the widening of the selectivity filter in TrkH. This is not wholly unexpected owing to

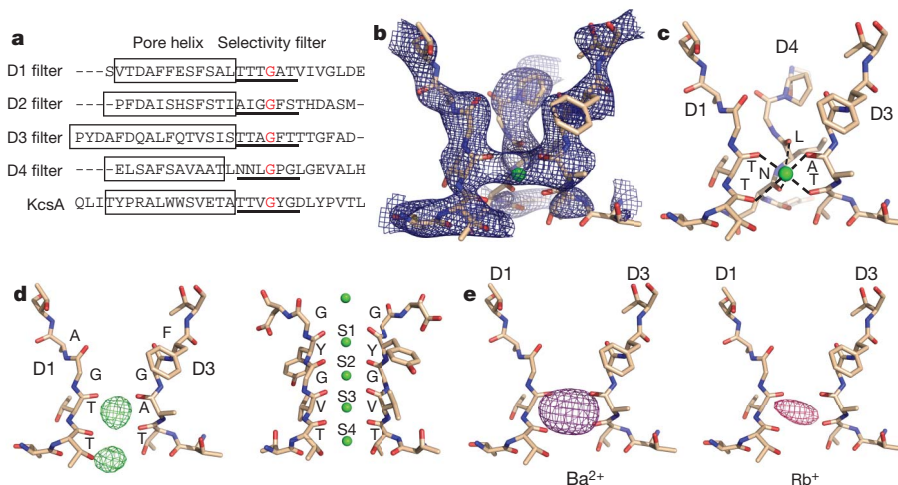


Figure 3 | Selectivity filter of VpTrkH. **a**, Amino-acid sequence alignment of the selectivity filter regions (underlined) and pore helices (box) in VpTrkH with the selectivity filter of the KcsA K^+ channel. The highly conserved glycine residue is marked in red. **b**, **c**, The selectivity filter with domain 2 removed, shown with **b** an NCS-averaged, simulated annealing omit map calculated with six residues from each selectivity filter omitted, contoured at 1σ , or **c** the coordination geometry of the K^+ (green sphere) highlighted. **d**, D1 and D3

from the K^+ structure are shown with $F_o - F_c$ electron density calculated without K^+ in the model and contoured at 3.5σ . The filter of KcsA is shown on the right for comparison. **e**, Ion-binding sites in the selectivity filter. Ba^{2+} (left) and Rb^+ (right) ($F_{o(ion)} - F_{o(K)}$) difference Fourier maps are shown contoured at 6.0 and 3.5σ levels, respectively, calculated using phases from the K^+ structure. The stick models are D1 and D3 from the K^+ structure.

the differences in the selectivity filter sequence between TrkH and KcsA: the second glycine residue in the signature sequence of K⁺ channels is not conserved in TrkH (Fig. 3a), and in the high-resolution KcsA structure³⁰ the backbone torsion angles of this residue lie in an unfavourable region of the Ramachandran plot for non-glycine residues. Regardless of the number of sites, the structure shows that at least one dehydrated K⁺ binds to VpTrkH through coordination by oxygen atoms in a manner similar to a K⁺ channel. Further experiments are needed to measure more accurately the occupancy of K⁺ and whether VpTrkH exhibits the high selectivity characteristic of K⁺ channels. In addition, because the four-site configuration is crucial for a high K⁺ throughput in K⁺ channels^{27,31}, we conjecture that VpTrkH, if it operates by a channel-like mechanism, conducts K⁺ with a lower throughput.

The ion permeation pathway

Although each protomer forms a continuous pore, a K⁺ permeating from the extracellular side encounters a barrier shortly after it exits the selectivity filter (Fig. 2c). The barrier is composed of two elements: Arg 468 from the second transmembrane segment of domain 4 (D4M2), and an intramembrane circular loop between D3M2a and D3M2b (Figs 2b and 4a). This is a unique feature and is not observed in any known K⁺ channel structures.

Arg 468 is conserved in almost all bacterial SKT proteins. The guanidinium group points towards the centre of the permeation pathway and is approximately 3.1 Å away from the backbone carbonyl oxygen atom of Gly 353, which is part of the intramembrane loop formed by Gly 346 to Lys 357 (Fig. 4a). This loop is present in all TrkH, TrkG and KtrB families, is rich in glycine and other small residues such as alanine and serine, and contains several highly conserved residues (Supplementary Figs 1 and 8). In addition, Glu 470,

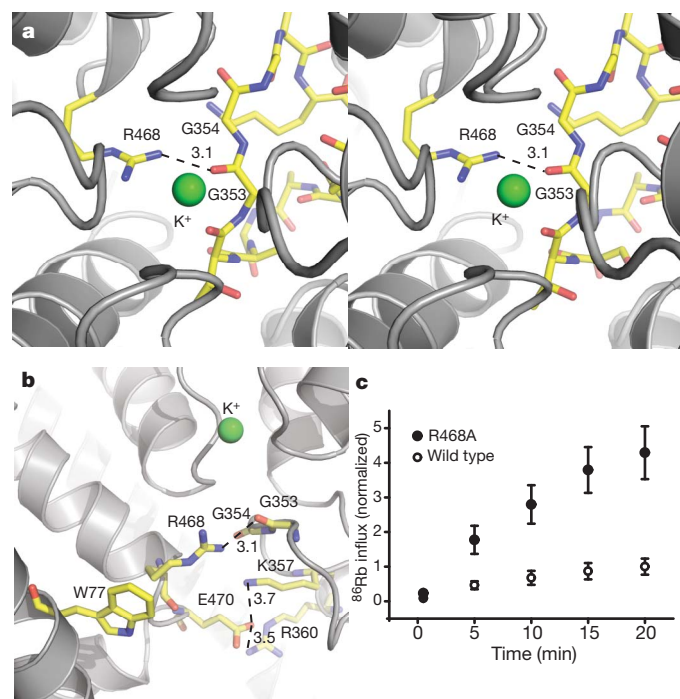


Figure 4 | Constriction formed by Arg 468 and the intramembrane loop. **a**, Stereo view of the interactions between the intramembrane loop and Arg 468 as viewed looking down the selectivity filter from the extracellular side. **b**, Interactions between the intramembrane loop and Arg 468 and Glu 470 as viewed from within the plane of the membrane. Residues Gly 353–354, Lys 357, Arg 360, Arg 468 and Glu 470 are shown as stick representations, and the dashed lines indicate distances between them. **c**, Time-dependent ⁸⁶Rb influx into proteoliposomes reconstituted with wild-type (open circle) or R468A (filled circle) VpTrkH. Error bars, s.e.m.

another highly conserved residue on D4M2, is close to Lys 357 in the loop and Arg 360 in D3M2b: a carboxylate oxygen on Glu 470 is 3.7 and 3.5 Å away from the terminal nitrogens of Lys 357 and Arg 360, respectively. These electrostatic interactions probably further stabilize the position of the intramembrane loop (Fig. 4b and Supplementary Fig. 9).

The narrow constriction formed by Arg 468 and the intramembrane loop has to widen for a K⁺ to reach the intracellular side. To understand further the role of Arg 468 in K⁺ permeation, we substituted it with an alanine and reconstituted the R468A mutant protein into liposomes. ⁸⁶Rb⁺ flux was then measured both for the wild type and for the R468A mutant in side-by-side experiments. Tracer uptake by the R468A mutant is significantly faster than that by the wild type (Fig. 4c), consistent with the observation that R468 occludes K⁺ permeation. An earlier study showed that mutating the equivalent of Arg 468 in a KtrB reduces K⁺ uptake in *E. coli*³², which seems to contradict our functional results and the role of Arg 468 inferred from the structure. However, because the KtrB mutant was assayed *in vivo*, additional factors such as expression levels or association with auxiliary proteins could have affected the measurement. As for the role of the intramembrane loop, a recent study showed that various deletions of the corresponding loop in a KtrB significantly increase K⁺ transport activity³³, consistent with its position shown by the structure.

Discussion

The pseudo-symmetry arising from duplication of an ancestral channel that is observed in the SKT family is not unique among ion channel/transporter families. In the animal kingdom, gene fusion or duplication of a more complex potassium channel, the voltage-dependent K⁺ channel (K_v), generated voltage-dependent Na⁺ and Ca²⁺ channels (Na_v and Ca_v, respectively)^{34,35}. Similar to the SKT proteins, four pseudo-subunits of K_v channels are expressed as a single polypeptide chain in the Na_v and Ca_v channels, although the domains are assembled in a clockwise direction when viewed from the extracellular side^{36,37} in contrast to the anticlockwise orientation observed in VpTrkH. The pseudo-fourfold symmetry in VpTrkH is strongest at the ion-binding sites in the selectivity filter region, where each domain contributes similarly to coordination of a K⁺. However, the symmetry starts to break down outside the selectivity filter where the pore helices vary in length and form different angles with the membrane norm, and the symmetry becomes considerably weaker for the M1 and M2 helices from different homologous domains. Especially notable is D3, which contains the intramembrane loop and the tilted D3M2b helix, and therefore is expected to contribute to gating of the permeation pathway more than the other domains. It is probable that in Na_v and Ca_v channels, the fourfold symmetry is maintained at the selectivity region for coordination of permeating ions but gradually breaks down so that the voltage-sensing modules and the connected gating machinery of a particular pseudo-subunit may contribute more than the others.

TrkH/TrkG and KtrB assemble with the cytosolic adenine nucleotide-binding proteins TrkA^{38,39} and KtrA⁴⁰, respectively. KtrA forms a ring and can undergo significant conformational changes upon binding to, or changes in the oxidation state of, ligands such as NAD and ATP^{20,41,42}. The ring has a twofold symmetry that matches that of the homodimeric assembly of the transmembrane subunits, and could potentially allosterically control K⁺ permeation. We speculate that the dimeric assembly is required for regulation of K⁺ transport by TrkA, although each protomer contains an independent ion permeation pathway. Although the structure of VpTrkH alone does not answer the question of whether TrkH operates as a channel or a transporter, it provides a framework for further studies that will reveal the molecular mechanism of K⁺ uptake and its regulation by the intracellular TrkA subunit.

METHODS SUMMARY

Both native and selenomethionine-substituted VpTrkH were expressed in *E. coli* with a C-terminal polyhistidine tag, extracted into decylmaltoside, and purified

by a metal-affinity column. The proteins were further purified by size-exclusion chromatography and concentrated to 8–10 mg ml⁻¹ for crystallization trials. Crystals were obtained by microbatch crystallization under paraffin oil. Co-crystallization with different ions was achieved by running the size-exclusion chromatography in buffers containing 150 mM of KCl or RbCl, or co-crystallization with BaCl₂. Initial phases were obtained by single-wavelength anomalous diffraction from a 3.9-Å data set collected on selenomethionine-derivatized VpTrkH, and a 3.5-Å native data set (with K⁺) was used for refinement of the final model. The final values of *R* and *R*_{free} were 24.9% and 29.9%, respectively. Purified VpTrkH was reconstituted into liposomes for ⁸⁶Rb⁺ flux assays as described previously^{17,22}.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 July; accepted 3 December 2010.

Published online 13 February 2011.

- Hille, B. *Ion Channels of Excitable Membranes* Ch. 1, 3rd edn (Sinauer, 2001).
- Schultz, S. G., Epstein, W. & Solomon, A. K. Cation transport in *Escherichia coli*. IV. Kinetics of net K uptake. *J. Gen. Physiol.* **47**, 329–346 (1963).
- Kroll, R. G. & Booth, I. R. The role of potassium transport in the generation of a pH gradient in *Escherichia coli*. *Biochem. J.* **198**, 691–698 (1981).
- Plack, R. H. Jr & Rosen, B. P. Cation/proton antiport systems in *Escherichia coli*. Absence of potassium/proton antiporter activity in a pH-sensitive mutant. *J. Biol. Chem.* **255**, 3824–3825 (1980).
- Bakker, E. P. & Mangerich, W. E. Interconversion of components of the bacterial proton motive force by electrogenic potassium transport. *J. Bacteriol.* **147**, 820–826 (1981).
- Skou, J. C. The influence of some cations on an adenosine triphosphatase from peripheral nerves. *Biochim. Biophys. Acta* **23**, 394–401 (1957).
- Tholema, N. et al. All four putative selectivity filter glycine residues in KtrB are essential for high affinity and selective K⁺ uptake by the KtrAB system from *Vibrio alginolyticus*. *J. Biol. Chem.* **280**, 41146–41154 (2005).
- Jan, L. Y. & Jan, Y. N. Cloned potassium channels from eukaryotes and prokaryotes. *Annu. Rev. Neurosci.* **20**, 91–123 (1997).
- Durell, S. R., Hao, Y., Nakamura, T., Bakker, E. P. & Guy, H. R. Evolutionary relationship between K⁺ channels and symporters. *Biophys. J.* **77**, 775–788 (1999).
- Durell, S. R. & Guy, H. R. Structural models of the KtrB, TrkH, and Trk1,2 symporters based on the structure of the KcsA K⁺ channel. *Biophys. J.* **77**, 789–807 (1999).
- Doyle, D. A. et al. The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity. *Science* **280**, 69–77 (1998).
- Durell, S. R., Bakker, E. P. & Guy, H. R. Does the KdpA subunit from the high affinity K⁺-translocating P-type KDP-ATase have a structure similar to that of K⁺ channels? *Biophys. J.* **78**, 188–199 (2000).
- Tholema, N., Bakker, E. P., Suzuki, A. & Nakamura, T. Change to alanine of one out of four selectivity filter glycines in KtrB causes a two orders of magnitude decrease in the affinities for both K⁺ and Na⁺ of the Na⁺ dependent K⁺ uptake system KtrAB from *Vibrio alginolyticus*. *FEBS Lett.* **450**, 217–220 (1999).
- Mäser, P. et al. Glycine residues in potassium channel-like selectivity filters determine potassium selectivity in four-loop-per-subunit HKT transporters from plants. *Proc. Natl Acad. Sci. USA* **99**, 6428–6433 (2002).
- Zeng, G. F., Pypaert, M. & Slayman, C. L. Epitope tagging of the yeast K⁺ carrier Trk2p demonstrates folding that is consistent with a channel-like structure. *J. Biol. Chem.* **279**, 3003–3013 (2004).
- Heginbotham, L., Lu, Z., Abramson, T. & MacKinnon, R. Mutations in the K⁺ channel signature sequence. *Biophys. J.* **66**, 1061–1067 (1994).
- Shi, N., Ye, S., Alam, A., Chen, L. & Jiang, Y. Atomic structure of a Na⁺- and K⁺-conducting channel. *Nature* **440**, 570–574 (2006).
- Stumpe, S. & Bakker, E. P. Requirement of a large K⁺-uptake capacity and of extracytoplasmic protease activity for protamine resistance of *Escherichia coli*. *Arch. Microbiol.* **167**, 126–136 (1997).
- Alkhuder, K., Meibom, K. L., Dubail, I., Dupuis, M. & Charbit, A. Identification of *trkH*, encoding a potassium uptake protein required for *Francisella tularensis* systemic dissemination in mice. *PLoS ONE* **5**, e8966 (2010).
- Albright, R. A., Ibar, J. L., Kim, C. U., Gruner, S. M. & Morais-Cabral, J. H. The RCK domain of the KtrAB K⁺ transporter: multiple conformations of an octameric ring. *Cell* **126**, 1147–1159 (2006).
- Albright, R. A., Joh, K. & Morais-Cabral, J. H. Probing the structure of the dimeric KtrB membrane protein. *J. Biol. Chem.* **282**, 35046–35055 (2007).
- Heginbotham, L., Kolmakova-Partensky, L. & Miller, C. Functional reconstitution of a prokaryotic K⁺ channel. *J. Gen. Physiol.* **111**, 741–749 (1998).
- Hendrickson, W. A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58 (1991).
- Daley, D. O. et al. Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308**, 1321–1323 (2005).
- Roux, B. & MacKinnon, R. The cavity and pore helices in the KcsA K⁺ channel: electrostatic stabilization of monovalent cations. *Science* **285**, 100–102 (1999).
- Lockless, S. W., Zhou, M. & MacKinnon, R. Structural and thermodynamic properties of selective ion binding in a K⁺ channel. *PLoS Biol.* **5**, e121 (2007).
- Morais-Cabral, J. H., Zhou, Y. & MacKinnon, R. Energetic optimization of ion conduction rate by the K⁺ selectivity filter. *Nature* **414**, 37–42 (2001).
- LeMasurier, M., Heginbotham, L. & Miller, C. KcsA: it's a potassium channel. *J. Gen. Physiol.* **118**, 303–314 (2001).
- Schlösser, A., Meldorf, M., Stumpe, S., Bakker, E. P. & Epstein, W. TrkH and its homolog, TrkG, determine the specificity and kinetics of cation transport by the Trk system of *Escherichia coli*. *J. Bacteriol.* **177**, 1908–1910 (1995).
- Zhou, Y., Morais-Cabral, J. H., Kaufman, A. & MacKinnon, R. Chemistry of ion coordination and hydration revealed by a K⁺ channel-Fab complex at 2.0 Å resolution. *Nature* **414**, 43–48 (2001).
- Zhou, M. & MacKinnon, R. A mutant KcsA K⁺ channel with altered conduction properties and selectivity filter ion distribution. *J. Mol. Biol.* **338**, 839–846 (2004).
- Kato, N. et al. Role of positively charged amino acids in the M2p transmembrane helix of Ktr/Trk/HKT type cation transporters. *Channels (Austin)* **1**, 161–171 (2007).
- Hänelt, I. et al. Gain of function mutations in membrane region M2C2 of KtrB open a gate controlling K⁺ transport by the KtrAB system from *Vibrio alginolyticus*. *J. Biol. Chem.* **285**, 10318–10327 (2010).
- Noda, M. et al. Primary structure of *Electrophorus electricus* sodium channel deduced from cDNA sequence. *Nature* **312**, 121–127 (1984).
- Mikami, A. et al. Primary structure and functional expression of the cardiac dihydropyridine-sensitive calcium channel. *Nature* **340**, 230–233 (1989).
- Li, R. A. et al. Clockwise domain arrangement of the sodium channel revealed by μ -conotoxin (GIIIA) docking orientation. *J. Biol. Chem.* **276**, 11072–11077 (2001).
- Stary, A., Shafir, Y., Hering, S., Wolschann, P. & Guy, H. R. Structural model of the Ca_v1.2 pore. *Channels (Austin)* **2**, 210–215 (2008).
- Bossemeyer, D. et al. K⁺-transport protein TrkA of *Escherichia coli* is a peripheral membrane protein that requires other *trk* gene products for attachment to the cytoplasmic membrane. *J. Biol. Chem.* **264**, 16403–16410 (1989).
- Schlösser, A., Hamann, A., Bossemeyer, D., Schneider, E. & Bakker, E. P. NAD⁺ binding to the *Escherichia coli* K⁺-uptake protein TrkA and sequence similarity between TrkA and domains of a family of dehydrogenases suggest a role for NAD⁺ in bacterial transport. *Mol. Microbiol.* **9**, 533–543 (1993).
- Nakamura, T., Yuda, R., Unemoto, T. & Bakker, E. P. KtrAB, a new type of bacterial K⁺-uptake system from *Vibrio alginolyticus*. *J. Bacteriol.* **180**, 3491–3494 (1998).
- Roosild, T. P., Miller, S., Booth, I. R. & Choe, S. A mechanism of regulating transmembrane potassium flux through a ligand-mediated conformational switch. *Cell* **109**, 781–791 (2002).
- Roosild, T. P. et al. KTN (RCK) domains regulate K⁺ channels and transporters by controlling the dimer-hinge conformation. *Structure* **17**, 893–903 (2009).
- Ho, B. K. & Gruswitz, F. HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct. Biol.* **8**, 49 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Data for this study were measured at beamlines X4A, X4C, X25 and X29 of the National Synchrotron Light Source and the NE-CAT 24ID-C and E at the Advanced Photon Source. This work was supported by the U.S. National Institutes of Health (grants HL086392, DK088057 and GM05026-sub0007 to M.Z.) and the American Heart Association (0630148N to M.Z.). M.Z. is a Pew Scholar in Biomedical Sciences. The New York Consortium on Membrane Protein Structure central facility is supported by grant GM05026 to W.A.H. as part of the Protein Structure Initiative (PSI-2) established by the National Institute of General Medical Sciences. We thank B. Honig for support, K. Jung for providing *E. coli* LB650, and J. Morais-Cabral, S.-Y. Lee, H. R. Guy, C. L. Slayman and E. P. Bakker for discussions and comments on the manuscript. M.Z. is grateful to R. MacKinnon for advice and support throughout the project.

Author Contributions M.P., J.L., B.R. and W.A.H. identified TrkH/TrkG/KtrB homologues in the database. R.B., B.K. and J.L. cloned and tested expression of the homologues. Y.C., H.H., J.W., E.J.L. and M.Z. scaled up production of proteins, produced and refined VpTrkH crystals, and collected and analysed X-ray diffraction data. X.J., E.J.L. and M.Z. solved and refined the structures. V.K., S.Y. and E.M.-H. analysed diffraction data and obtained a partial model in early stages of the project. Y.C., M.G.D., M.Q., Y.P., Y.J., J.A.J. and M.Z. characterized VpTrkH function. K.R.R. and W.A.H. advised on data collection and crystallography. E.J.L. and M.Z. wrote the manuscript with inputs from all authors.

Author Information Atomic coordinates and structure factors are deposited in Protein Data Bank under accession number 3PJZ. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.Z. (mz2140@columbia.edu).

METHODS

Homology screen, cloning and initial protein expression. TrkH was first established to be a valid target for structural studies by a bioinformatics analysis^{44,45}. A total of 91 *trkG/trkH/ktrB* genes from 58 prokaryotic genomes were identified, and the genes were amplified by PCR from the genomic DNAs, inserted into a modified pET plasmid (Novagen) with a C-terminal deca-histidine tag and a TEV protease recognition site, and expressed in a small-scale culture. Protein expression was then examined using western blots as a readout, and western-positive clones were pursued for further study. Identification and cloning of homologues, and the initial expression study, were performed by a high-throughput approach in the central facility of the New York Consortium on Membrane Protein Structure; a detailed description of the procedures can be found in ref. 44.

Large-scale protein expression, purification and crystallization. Western-positive clones received from the New York Consortium on Membrane Protein Structure were scaled up for mid- to large-scale expression studies. Five proteins (TrkHs from *V. parahaemolyticus*, *Vibrio fischeri*, *Idiomarina loihiensis*, *Campylobacter jejuni* and a KtrB from *V. fischeri*) had yields higher than 0.25 mg l⁻¹ cell culture and three of them (TrkHs from *V. parahaemolyticus*, *I. loihiensis* and *C. jejuni*) exhibited a mono-dispersed profile in size-exclusion chromatography. Among those proteins, only TrkH from *V. parahaemolyticus* (VpTrkH) produced diffracting crystals and thus became the focus of crystallization efforts.

For large-scale purification of native VpTrkH, BL21(DE3) cells were grown in Luria broth at 37 °C and induced with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) after absorbance ($A_{600\text{ nm}}$) reached 1.0; for expression of selenomethionine-incorporated proteins, the cells were grown in minimal medium containing 32.2 mM K₂HPO₄, 11.7 mM KH₂PO₄, 6 mM (NH₄)₂SO₄, 0.68 mM Na citrate, 0.17 mM Mg₂SO₄, 32 mM glucose, 0.008% (w/v) alanine, arginine, aspartic acid, asparagine, cysteine, glutamic acid, glycine, histidine, proline, serine, tryptophan, glutamine and tyrosine, 0.02% (w/v) isoleucine, leucine, lysine, phenylalanine, threonine and valine, 25 mg l⁻¹ L-selenium-methionine, 32 mg l⁻¹ thiamine, and 32 mg l⁻¹ thymine, and induced at an absorbance ($A_{600\text{ nm}}$) of 0.6. The cell membranes were solubilized with 40 mM *n*-decyl- β -D-maltoside (Anatrace) and the His-tagged protein was purified with TALON Metal Affinity Resin (Clontech). After removal of the tag with tobacco etch virus protease, the native protein was subjected to size-exclusion chromatography with a Superdex 200 10/300 GL column (GE Health Sciences) equilibrated in a buffer of 150 mM KCl, 20 mM HEPES, pH 7.5, 5 mM β -mercaptoethanol and 3.5 mM *n*-decyl- β -D-maltoside. The selenomethionine-incorporated VpTrkH protein was purified by the same procedure. The native protein was concentrated to 8 mg ml⁻¹ and the selenomethionine-substituted protein to 10 mg ml⁻¹ as approximated by ultraviolet absorbance.

Although both the native and selenomethionine-substituted VpTrkH yielded crystals readily, the crystals diffracted anisotropically and an overwhelming majority failed to reach 4.0 Å. More than 3,000 crystals were screened over a period of over 3 years. Native VpTrkH crystals were grown by microbatch crystallization under paraffin oil where 1.5 μ l of the protein solution was mixed with an equal volume of crystallization solution containing 35% PEG400, 200 mM calcium acetate and 100 mM sodium acetate, pH 5.3. Rb⁺-derivatized crystals were obtained by the same method except that the size-exclusion chromatography during purification was conducted in a buffer using 150 mM RbCl to replace KCl. Ba²⁺-derivatized crystals were obtained by adding 10 mM BaCl₂ into the native protein before mixing with crystallization solution. Before flash-freezing in liquid nitrogen, the crystals were cryoprotected in mother liquor for 2–5 s. The mother liquor was obtained by vapour diffusion in sitting drops mixed from 3 μ l of the protein solution and an equal volume of well solution containing 35% PEG400, 200 mM calcium acetate and 100 mM sodium acetate, pH 5.3.

Data collection and structure solution. Diffraction data were collected on beamlines X25 and X29 at the National Synchrotron Light Source and 24ID-C and 24ID-E at the Advanced Photon Source. A 3.9-Å data set was collected at a wavelength of 0.9791 Å on selenomethionine-derivatized VpTrkH. The data were indexed, integrated and scaled using HKL2000 (ref. 46). The data set showed anisotropy, but nonetheless exhibited a strong anomalous signal and was therefore used to obtain the initial phases to 3.9 Å by single-wavelength anomalous diffraction. The positions of 18 out of 24 available selenium sites in the asymmetric unit were located using the program phenix.hyss⁴⁷. Initial experimental phases were improved by twofold NCS averaging and solvent flattening using RESOLVE⁴⁷. The resultant density-modified experimental map was used to build manually a partial C α trace with COOT⁴⁸. The phases were gradually extended to a higher-resolution native data set at 3.5 Å using twofold NCS averaging, solvent flattening and histogram matching in DM⁴⁹. Manual model building and sequence assignments were done iteratively using COOT, and the structure

refinement was done using PHENIX⁴⁷ with strong twofold NCS restraints. The final refined model contained residues 1–157, 174–484, and one K⁺ in each subunit in the asymmetric unit. The anomalous difference Fourier map calculated with phases from the refined model confirmed the correctness of the initial selenium sites, all of which overlay well with ordered methionine residues in the final model, and identified two additional sites corresponding to two N-terminal methionines in the asymmetric unit (Supplementary Fig. 10a–c). The region encompassing residues 158–173 is disordered, consistent with the absence of anomalous peaks expected from Met 158 and Met 174 within this region. At the model-building stage, diffraction data were corrected for anisotropy using the Anisotropy Server⁵⁰ and a second model was refined with anisotropy correction. Models with or without the correction essentially overlap; however, map quality is improved in several regions after the correction.

The final refined model devoid of K⁺ was used to calculate the $F_{\text{Rb}} - F_{\text{K}}$ and $F_{\text{Ba}} - F_{\text{K}}$ difference maps for structures containing Rb⁺ and Ba²⁺.

E. coli complementation assay. *E. coli* strain LB650 competent cells were transformed with pET31 vector carrying *V. parahaemolyticus trkH*. Two media, Hi-K and Lo-K, were prepared based on ref. 51, both containing 8 mM ammonium sulphate, 0.4 mM magnesium sulphate, 1 mM sodium citrate, 1 mg l⁻¹ thiamine and 2 g l⁻¹ glucose. For the Hi-K medium, 115 mM potassium phosphate (pH 7.0) was added; for the Lo-K medium, 115 mM sodium phosphate (pH 7.0) was added. The two solutions were mixed in different ratios to produce the desired K⁺ concentration. The transformation cell culture was spread onto agar plates prepared with solutions with different K⁺ concentrations and incubated at 37 °C overnight. As a blank control, pET31 vector without insertions was used to transform *E. coli* strain LB650 competent cells.

Reconstitution of TrkH into proteoliposomes. Purified VpTrkH was reconstituted into lipid vesicles composed of 1-palmitoyl-2-oleoyl-phosphatidylethanolamine and 1-palmitoyl-2-oleoyl-phosphatidylglycerol (Avanti Polar Lipids) in a ratio of 3:1 by mass, as previously described⁵¹. The protein to lipid ratio was 1:150 by mass. The solution was then subjected to several rounds of dialysis against reconstitution solution until the detergent was totally removed. At the end of each experiment, valinomycin was added to the reaction mixture to monitor the maximum uptake.

Determination of protein oligomeric state. The mass of the VpTrkH protein in solution was measured using a Wyatt miniDAWN TREOS 3 angle-static light-scattering detector, a Wyatt Optilab rEX refractive index detector and an Agilent Variable Wavelength Detector ultraviolet absorbance detector⁵². Purified protein sample (5 μ l) was injected onto a TSKgel SuperSW3000 (4.6 mm internal diameter \times 30 cm) silica-gel size-exclusion column in buffer containing 0.016% dodecyl-maltoside at a rate of 0.25 ml min⁻¹. The differential refractive index (dn/dc) value for dodecylmaltoside was calculated to be 0.128 ml g⁻¹ using the Wyatt Optilab rEX refractive index detector. Deconvolution of the protein–detergent conjugate was then achieved using a previously described method⁵³. The calculation did not account for refractive index contributions due to bound lipid.

⁸⁶Rb flux assay. The ⁸⁶Rb⁺ flux assay was performed as described previously²². In the competition assays, Li⁺, Na⁺, K⁺ or Rb⁺ were added directly into the flux buffer, and the readings were taken at the 20 min time point.

44. Love, J. *et al.* The New York Consortium on Membrane Protein Structure (NYCOMPS): a high-throughput platform for structural genomics of integral membrane proteins. *J. Struct. Funct. Genomics* **11**, 191–199 (2010).
45. Punta, M. *et al.* Structural genomics target selection for the New York consortium on membrane protein structure. *J. Struct. Funct. Genomics* **10**, 255–268 (2009).
46. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
47. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
48. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
49. Cowtan, K. Error estimation and bias correction in phase-improvement calculations. *Acta Crystallogr. D* **55**, 1555–1567 (1999).
50. Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103**, 8060–8065 (2006).
51. Epstein, W. & Kim, B. S. Potassium transport loci in *Escherichia coli* K-12. *J. Bacteriol.* **108**, 639–644 (1971).
52. Slotboom, D. J., Duurkens, R. H., Olieman, K. & Erkens, G. B. Static light scattering to characterize membrane proteins in detergent solution. *Methods* **46**, 73–82 (2008).
53. Kendrick, B. S., Kerwin, B. A., Chang, B. S. & Philo, J. S. Online size-exclusion high-performance liquid chromatography light scattering and differential refractometry methods to determine degree of polymer conjugation to proteins and protein–protein or protein–ligand association states. *Anal. Biochem.* **299**, 136–146 (2001).

An hour-glass magnetic spectrum in an insulating, hole-doped antiferromagnet

A. T. Boothroyd¹, P. Babkevich^{1,2}, D. Prabhakaran¹ & P. G. Freeman^{3†}

Superconductivity in layered copper oxide compounds emerges when charge carriers are added to antiferromagnetically ordered CuO_2 layers¹. The carriers destroy the antiferromagnetic order, but strong spin fluctuations persist throughout the superconducting phase and are intimately linked to superconductivity². Neutron scattering measurements of spin fluctuations in hole-doped copper oxides have revealed an unusual ‘hour-glass’ feature in the momentum-resolved magnetic spectrum that is present in a wide range of superconducting and non-superconducting materials^{3–15}. There is no widely accepted explanation for this feature. One possibility is that it derives from a pattern of alternating spin and charge stripes¹⁶, and this idea is supported by measurements on stripe-ordered $\text{La}_{1.875}\text{Ba}_{0.125}\text{CuO}_4$ (ref. 15). Many copper oxides without stripe order, however, also exhibit an hour-glass spectrum^{3–12}. Here we report the observation of an hour-glass magnetic spectrum in a hole-doped antiferromagnet from outside the family of superconducting copper oxides. Our system has stripe correlations and is an insulator, which means that its magnetic dynamics can conclusively be ascribed to stripes. The results provide compelling evidence that the hour-glass spectrum in the copper oxide superconductors arises from fluctuating stripes.

The term ‘hour-glass’ describes the general structure of the magnetic spectrum as a function of energy E and wavevector \mathbf{Q} . At low energies, there is a four-fold pattern of incommensurate peaks centred on the antiferromagnetic wavevector of the parent (undoped) CuO_2 square lattice. The peaks disperse inwards with increasing energy until they meet at the antiferromagnetic wavevector, then disperse outwards again. A square-shaped intensity distribution is observed above the meeting point and is rotated by 45° with respect to that below the meeting point. The hour-glass shape is common to many if not all hole-doped copper oxides, and is especially prominent in underdoped compositions. It is observed regardless of whether the quartet of low-energy peaks is aligned along the Cu–O bonds or is at 45° to them, as found in lightly doped $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ (ref. 13). However, owing to the influence of superconductivity on the spin fluctuations, there are some important system- and temperature-dependent variations at low energies¹⁷.

Our experiment was performed on a member of the $\text{La}_{2-x}\text{Sr}_x\text{CoO}_4$ (LSCoO) series, which is isostructural with the ‘214’ copper oxide family (for example, $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$). The crystal structure of LSCoO contains near-perfect square lattices of Co atoms on well-separated CoO_2 layers (Fig. 1a). Unlike the copper oxides, LSCoO remains insulating over a wide range of doping¹⁸. The parent phase La_2CoO_4 exhibits¹⁹ commensurate antiferromagnetic order below the Néel temperature $T_N = 275$ K, in a fashion similar to La_2CuO_4 . Substitution of Sr for La donates positive charge onto the CoO_2 layers. When $x > 0.3$, the antiferromagnetic order of the parent phase is modulated at 45° to the Co–O bonds²⁰. This is attributed to a self-organization of holes into arrays of charged stripes, which create antiphase domain walls in the antiferromagnetic order. The existence of this stripe-like order is well established in the isostructural hole-doped nickel oxides $\text{La}_{2-x}\text{Sr}_x\text{NiO}_{4+y}$ (refs 21–23) and in the 214 copper oxides with $x \approx 1/8$ (ref. 16).

We studied LSCoO with $x = 1/3$. The ideal striped pattern of magnetic and charge order for $x = 1/3$ (Fig. 1b) consists of diagonal lines of Co^{3+} , three lattice spacings apart, separating bands of antiferromagnetically ordered Co^{2+} . The Co^{2+} ions are in the high-spin $S = 3/2$ state¹⁹ while the Co^{3+} are in the low-spin $S = 0$ state and so do not carry a moment²⁴ (see also Supplementary Information). The diagonal modulation of the antiferromagnetic order gives rise to magnetic diffraction peaks at wavevectors $\mathbf{Q}_m = \mathbf{Q}_{\text{AF}} \pm (1/6, 1/6, 0)$ and equivalent positions in reciprocal space (Fig. 1c), where $\mathbf{Q}_{\text{AF}} = (h + 1/2, k + 1/2, l)$ is the antiferromagnetic (AF) wavevector and h, k and l are integers. The antiferromagnetic order can equally well be modulated along the other diagonal, giving peaks at $\mathbf{Q}_m = \mathbf{Q}_{\text{AF}} \pm (1/6, -1/6, 0)$. In reality, peaks from both orthogonal domains are present with equal intensity.

Figure 2 is a composite image of the measured magnetic spectrum for wavevectors in the $(Q_x, Q_y, 0)$ plane in reciprocal space. The elastic ($E = 0$ meV) scattering contains four magnetic peaks centred on the \mathbf{Q}_m positions. The peaks are significantly broader than the instrumental resolution and are elongated in the direction perpendicular to the stripes. From the measured half-widths, we determined correlation

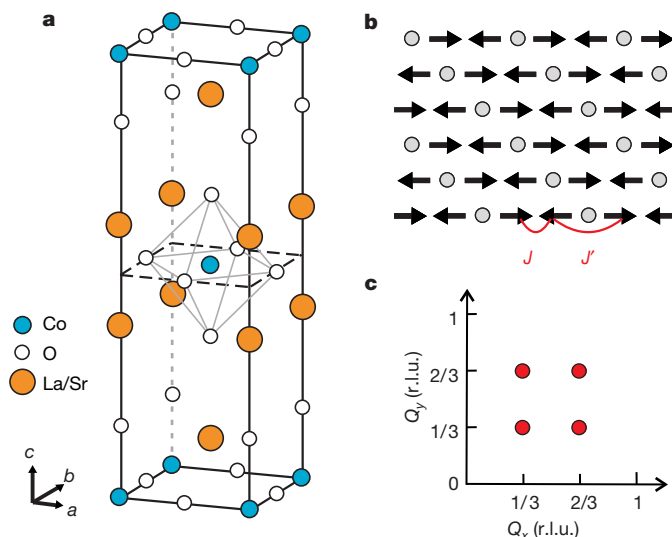


Figure 1 | Crystal structure and magnetic order of $\text{La}_{5/3}\text{Sr}_{1/3}\text{CoO}_4$. **a**, Conventional unit cell of the tetragonal $I4/mmm$ structure. The cell parameters are $a = b = 3.87$ Å, $c = 12.6$ Å. **b**, Ideal model of magnetic and charge order in a CoO_2 layer. The arrows represent ordered moments on Co^{2+} and the circles represent non-magnetic Co^{3+} . Oxygen atoms are omitted for clarity. The intra- and inter-stripe exchange interactions (respectively J and J') are shown. **c**, Diagram of reciprocal space showing the magnetic peak positions (filled circles) projected onto the (Q_x, Q_y) plane. The magnetic peaks at $(1/3, 1/3)$ and $(2/3, 2/3)$ correspond to the stripe domain in **b** while those at $(1/3, 2/3)$ and $(2/3, 1/3)$ are from the equivalent domain in which the stripes run along the opposite diagonal. Wavevectors are given in reciprocal lattice units (r.l.u.) relative to the tetragonal cell.

¹Department of Physics, University of Oxford, Clarendon Laboratory, Oxford OX1 3PU, UK. ²Laboratory for Neutron Scattering, Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland. ³Institut Laue-Langevin, BP 156, 38042 Grenoble Cedex 9, France. [†]Present address: Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner Platz 1, D-14109 Berlin, Germany.

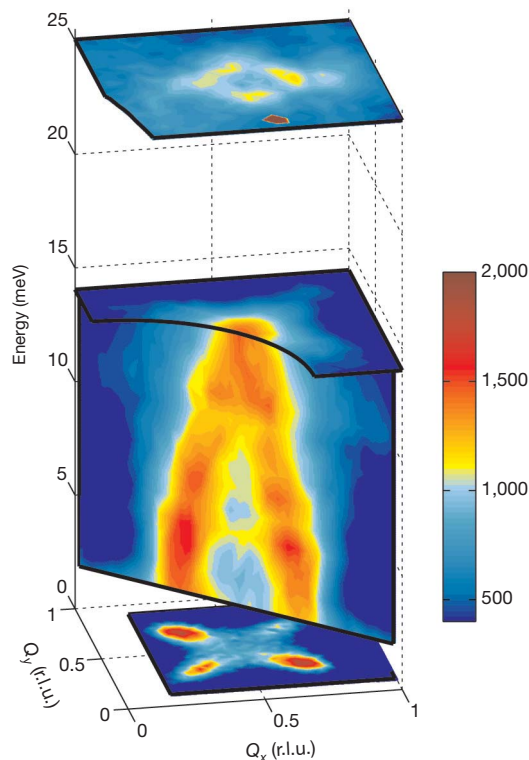


Figure 2 | Neutron scattering intensity maps of the magnetic excitation spectrum of $\text{La}_{5/3}\text{Sr}_{1/3}\text{CoO}_4$. The elastic and 14 meV maps (respectively bottom and centre) are centred on (0.5, 0.5, 0), and the 25 meV map (top) is centred on (1.5, 0.5, 0). The energy–wavevector map was constructed from a series of constant-energy scans made at 1 meV intervals through (0.5, 0.5, 0) in the $(\xi, -\xi, 0)$ direction. All measurements were made on a single crystal (Supplementary Information) at a temperature of 2 K, but a background recorded at 100 K has been subtracted from the elastic map. The colour scale shows neutron counts for the 14 meV and energy–wavevector maps. The elastic and 25 meV maps have been scaled to facilitate comparison on the same colour scale. The sharp feature in the 25 meV map near (1.5, 0, 0) is spurious. The data here and in Figs 3 and 4 were recorded on the IN8 triple-axis spectrometer at the Institut Laue–Langevin with a fixed final energy of either 14.7 meV or 34.8 meV, set by Bragg reflection from a graphite analyser. The incident energy was selected by Bragg reflection from a silicon ($E < 35$ meV) or copper ($E \geq 35$ meV) monochromator. A graphite filter was placed after the sample to suppress contamination from higher orders. No collimation was used. The sample was mounted in a helium cryostat and aligned with the a and b axes in the horizontal scattering plane.

lengths of $\xi_{\parallel} \approx 10 \text{ \AA}$ and $\xi_{\perp} \approx 6.5 \text{ \AA}$ parallel and perpendicular to the stripes, respectively. The magnetic peaks are observed below a temperature $T_N \approx 100 \text{ K}$ (Supplementary Fig. 1).

The inelastic response in Fig. 2 begins at low energies with ‘legs’ of scattering which rise from the \mathbf{Q}_m positions. The legs disperse inwards with increasing energy and meet at about 14 meV. At this energy, the intensity in the $(Q_x, Q_y, 0)$ plane is peaked at \mathbf{Q}_{AF} . The peak has an anisotropic shape with protrusions along the diagonals of the reciprocal lattice. The intensity pattern shown at 25 meV retains four-fold symmetry but is rotated 45° with respect to the pattern below 14 meV.

The structure of the magnetic spectrum can be seen in more detail in Figs 3 and 4a–d. The constant-energy scans in Fig. 3a show that the inward dispersion of the magnetic peaks at low energies is not accompanied by a corresponding outward-dispersing feature, as would normally be expected for spin waves. Figure 3a and b shows that above ~ 20 meV the maximum intensity disperses away from \mathbf{Q}_{AF} . Between 14 meV and 20 meV, the intensity remains peaked at \mathbf{Q}_{AF} . The positions of peaks in these and similar scans are plotted in Fig. 3c, revealing a dispersion in the shape of an hour-glass. At higher energies, the signal converges to the antiferromagnetic zone corners, where the scattering

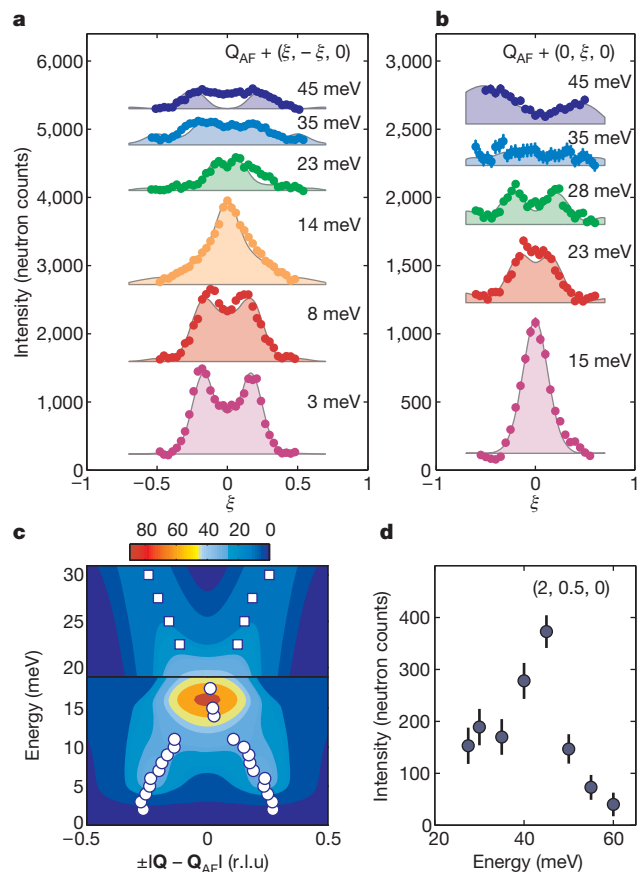


Figure 3 | Dispersion of the magnetic excitation spectrum of $\text{La}_{5/3}\text{Sr}_{1/3}\text{CoO}_4$. **a, b,** The variation in the scattered intensity with wavevector in diagonal and vertical scans through \mathbf{Q}_{AF} for a series of different excitation energies. We used $\mathbf{Q}_{AF} = (0.5, 0.5, 0)$ for $E < 15$ meV and $\mathbf{Q}_{AF} = (1.5, 0.5, 0)$ for $E \geq 15$ meV. Successive scans have been displaced vertically for clarity, and the intensities at 35 meV and 45 meV have been scaled to facilitate comparison with the lower-energy data. The shaded peaks are simulations from the spin-wave model (see Fig. 4). The intensity of the calculated signal was adjusted to match the data, and a linear background was added. **c,** Dispersion of the intensity in the magnetic spectrum. The symbols represent the centres of Gaussian or Lorentzian peaks fitted to those constant-energy scans which show either two clearly resolved peaks or a single central peak, circles from scans parallel to $(\xi, -\xi, 0)$ and squares from scans parallel to $(0, \xi, 0)$. The colour map is a simulation of the neutron scattering intensity from the spin-wave model, colour-coded in arbitrary units. For energies below 19 meV the wavevector axis represents a diagonal scan through \mathbf{Q}_{AF} , and for energies above 19 meV it is a vertical scan. **d,** Energy dependence of the magnetic intensity at the point (2, 0, 0). A background measured at (2, 0, 0) has been subtracted. Error bars, s.d. The strong peak at ~ 45 meV corresponds to the top of the magnetic spectrum. All measurements were made at a temperature of 2 K.

is strongly enhanced (Fig. 4d). Figure 3d is the magnetic signal at one such point, and shows that the spectrum extends up to an energy of ~ 45 meV.

What is remarkable about these data is how closely they resemble the magnetic spectrum of the hole-doped copper oxides. Not only does the dispersion of the magnetic peaks have the characteristic hour-glass shape (Fig. 3c), but also there is (1) the same apparent absence of an outward-dispersing excitation branch emerging from \mathbf{Q}_m (Fig. 3a), and (2) the same 45° rotation of the four-fold intensity pattern on crossing from below to above the waist of the hour-glass (Figs 2 and 4a–c).

As LSCoO is an insulator, a local-moment description applies. Our analysis is based on the spin-wave spectrum of the ideal period-3 stripe pattern, Fig. 1b. A two-dimensional model is sufficient because the inter-layer coupling is weak (Supplementary Fig. 2). The strong planar

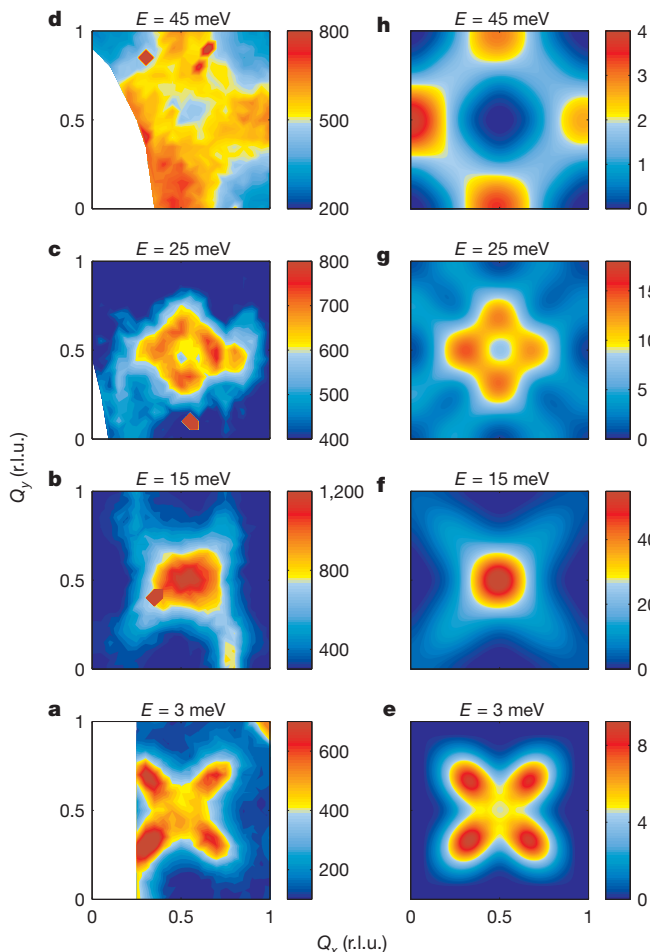


Figure 4 | Simulations of the magnetic spectrum of $\text{La}_{5/3}\text{Sr}_{1/3}\text{CoO}_4$. **a–d**, Constant-energy slices through the inelastic neutron scattering data. The regions of missing data at small \mathbf{Q} in the 3, 25 and 45 meV maps are due to experimental limits on the scattering angle. The colour scale shows neutron counts, but different counting times were used for each map. Sharp spots visible on the 15, 25 and 45 meV maps are spurious. **e–h**, The corresponding simulations from the spin-wave model described in ref. 25 and in Supplementary Information. The best-fit values of the exchange parameters (see Fig. 1b) were $J = 11.5$ meV and $J' = 0.55$ meV. The fixed parameters of the model are given in Supplementary Information. The simulated spectra were averaged over an equal population of four equivalent magnetic domains, two wavevector domains to allow for the alternative diagonals along which the stripes may run, and two spin domains to allow the ordered moments to align along the x and y axes with equal probability. To simulate the observed broadening, the δ -function in energy in the spin-wave cross-section was replaced by a Gaussian with a standard deviation of 1 meV and the spectrum further broadened by convolution with a two-dimensional Gaussian in wavevector with standard deviations 0.06 r.l.u. and 0.09 r.l.u. parallel and perpendicular to the stripes, respectively. The model includes the \mathbf{Q} variation of the dipole magnetic form factor of Co^{2+} . A common intensity scale (in arbitrary units) is used for **e–h**, indicated by the colour coding.

single-ion anisotropy of Co^{2+} is included in the model²⁵, and the two principal exchange parameters J and J' (Fig. 1b) were adjusted together with an intensity scale factor to fit the data.

Intensity calculations from the model are included in Figs 3 and 4. All the prominent features of the data are reproduced by the simulations. In particular, a good description of the hour-glass dispersion (Fig. 3c) and of the intensity distribution in the Q_x – Q_y plane (Fig. 4) is obtained. Moreover, the fitted values of J and J' are comparable with separate estimates of these parameters from LSCO with $x = 0$ and $x = 0.5$, respectively^{25,26} (Supplementary Information). These findings are strong evidence that the basis of our model—that is, the nature of the ground state and all the important interactions—is correct.

Analysis of the model reveals that the apparent inwards dispersion at low energies is caused by the combined effects of broadening and a larger intensity on the surface of the spin-wave cone nearest \mathbf{Q}_{AF} due to the high J/J' ratio²⁷ (this analysis also explains the absence of an hour-glass feature in the nickel oxides, see Supplementary Information). The point where the inward-dispersing branches meet is a saddle point formed by a maximum in the inter-stripe dispersion and a minimum in the intra-stripe dispersion. The superposition of dispersion surfaces from orthogonal stripe domains creates four intensity maxima above the saddle point, which are rotated 45° with respect to the magnetic peaks below the saddle point. The enhanced intensity near the top of the spectrum is due to the planar anisotropy in LSCO which creates a nearly flat dispersion at the highest energies (Supplementary Information).

Although our model assumes static magnetic order, the results are also relevant for slowly fluctuating stripes. This is because neutron scattering is insensitive to fluctuations much slower than $\sim \hbar/\Delta E \approx 10^{-11}$ s; here $\Delta E \approx 1$ meV is the energy resolution and \hbar is Planck's constant. A system in which the order parameter fluctuates more slowly than this has nearly the same spectrum as one in which the correlations are static.

The observation of an hour-glass magnetic spectrum in LSCO ($x = 1/3$) is significant because it provides an experimental demonstration that the hour-glass spectrum can arise from a system of slowly fluctuating magnetic stripes. This could have important implications for the copper oxide superconductors. The striking resemblance between the magnetic spectra of the layered copper oxides and that found here (apart from superconductivity-induced effects) suggests that the magnetic fluctuations in the copper oxides have the same stripe-like characteristics as those that cause the hour-glass spectrum in LSCO ($x = 1/3$), namely, (1) unidirectionally modulated anti-ferromagnetic correlations, and (2) a large ratio of the magnetic couplings parallel and perpendicular to the stripes. A further requirement is a degree of broadening in order to smear out the spin-wave-like dispersion cones. Although the spins can be very well correlated in the copper oxides at low energies, experiments indicate that the spectrum broadens with increasing energy¹¹.

The general conditions for an hour-glass magnetic spectrum found here could in principle be satisfied by different classes of microscopic model. They do not imply that the local-moment picture that describes the cobalt oxides necessarily extends to the copper oxides. They do, however, impose significant constraints on any such model. It is likely that magnetic stripe correlations in the copper oxides would be accompanied by charge-stripe correlations²⁸. Therefore, our results lend support to interpretations of the hour-glass spectrum based on disordered stripes, and provide encouragement for theories of the copper oxide superconductors in which the doped holes form a state with slowly fluctuating electronic nematic order^{28,29}.

Received 21 October 2010; accepted 3 February 2011.

1. Lee, P. A., Nagaosa, N. & Wen, X.-G. Doping a Mott insulator: physics of high-temperature superconductivity. *Rev. Mod. Phys.* **78**, 17–85 (2006).
2. Chubukov, A. V., Pines, D. & Schmalien, J. in *The Physics of Superconductors* Vol. 1, *Conventional and High- T_c Superconductors* (eds Bennemann, K. H. & Ketterson, J. B.) 495–590 (Springer, 2003).
3. Arai, M. *et al.* Incommensurate spin dynamics of underdoped superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6.7}$. *Phys. Rev. Lett.* **83**, 608–611 (1999).
4. Bourges, P. *et al.* The spin excitation spectrum in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{6.85}$. *Science* **288**, 1234–1237 (2000).
5. Hayden, S. M., Mook, H. A., Dai, P., Perring, T. G. & Doğan, F. The structure of the high-energy spin excitations in a high-transition-temperature superconductor. *Nature* **429**, 531–534 (2004).
6. Reznik, D. *et al.* Dispersion of magnetic excitations in optimally doped superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{6.95}$. *Phys. Rev. Lett.* **93**, 207003 (2004).
7. Stock, C. *et al.* From incommensurate to dispersive spin-fluctuations: the high-energy inelastic spectrum in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$. *Phys. Rev. B* **71**, 024522 (2005).
8. Hinkov, V. *et al.* Spin dynamics in the pseudogap state of a high-temperature superconductor. *Nature Phys.* **3**, 780–785 (2007).
9. Fauqué, B. *et al.* Dispersion of the odd magnetic resonant mode in near-optimally doped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Phys. Rev. B* **76**, 214512 (2007).

10. Xu, G. *et al.* Testing the itinerancy of spin dynamics in superconducting $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Nature Phys.* **5**, 642–646 (2009).
11. Christensen, N. B. *et al.* Dispersive excitations in the high-temperature superconductor $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Phys. Rev. Lett.* **93**, 147002 (2004).
12. Vignolle, B. *et al.* Two energy scales in the spin excitations of the high-temperature superconductor $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Nature Phys.* **3**, 163–167 (2007).
13. Matsuda, M. *et al.* Magnetic dispersion of the diagonal incommensurate phase in lightly doped $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Phys. Rev. Lett.* **101**, 197001 (2008).
14. Lipscombe, O. J., Vignolle, B., Perring, T. G., Frost, C. D. & Hayden, S. M. Emergence of coherent magnetic excitations in the high temperature underdoped $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ superconductor at low temperatures. *Phys. Rev. Lett.* **102**, 167002 (2009).
15. Tranquada, J. M. *et al.* Quantum magnetic excitations from stripes in copper oxide superconductors. *Nature* **429**, 534–538 (2004).
16. Tranquada, J. M., Sternlieb, B. J., Axe, J. D., Nakamura, Y. & Uchida, S. Evidence for stripe correlations of spins and holes in copper oxide superconductors. *Nature* **375**, 561–563 (1995).
17. Hayden, S. M. in *Superconductivity: Conventional and Unconventional Superconductors* Vol. 2 (eds Bennemann, K. H. & Ketterson, J. B.) 993–1029 (Springer, 2008).
18. Moritomo, Y., Higashi, K., Matsuda, K. & Nakamura, A. Spin-state transition in layered perovskite cobalt oxides: $\text{La}_{2-x}\text{Sr}_x\text{CoO}_4$ ($0.4 \leq x \leq 1.0$). *Phys. Rev. B* **55**, R14725–R14728 (1997).
19. Yamada, K. *et al.* Successive antiferromagnetic phase transitions in single-crystal La_2CoO_4 . *Phys. Rev. B* **39**, 2336–2343 (1989).
20. Cwik, M. *et al.* Magnetic correlations in $\text{La}_{2-x}\text{Sr}_x\text{CoO}_4$ studied by neutron scattering: possible evidence for stripe phases. *Phys. Rev. Lett.* **102**, 057201 (2009).
21. Chen, C. H., Cheong, S.-W. & Cooper, A. S. Charge modulations in $\text{La}_{2-x}\text{Sr}_x\text{NiO}_{4+y}$: ordering of polarons. *Phys. Rev. Lett.* **71**, 2461–2464 (1993).
22. Tranquada, J. M., Buttrey, D. J., Sachan, V. & Lorenzo, J. E. Simultaneous ordering of holes and spin in $\text{La}_2\text{NiO}_{4.125}$. *Phys. Rev. Lett.* **73**, 1003–1006 (1994).
23. Yoshizawa, H. *et al.* Stripe order at low temperatures in $\text{La}_{2-x}\text{Sr}_x\text{NiO}_4$ with $0.289 \leq x \leq 0.5$. *Phys. Rev. B* **61**, R854–R857 (2000).
24. Hollmann, N. *et al.* Anisotropic susceptibility of $\text{La}_{2-x}\text{Sr}_x\text{CoO}_4$ related to the spin states of cobalt. *N. J. Phys.* **10**, 023018 (2008).
25. Helme, L. M. *et al.* Magnetic order and dynamics of the charge-ordered antiferromagnet $\text{La}_{1.5}\text{Sr}_{0.5}\text{CoO}_4$. *Phys. Rev. B* **80**, 134414 (2009).
26. Babkevich, P., Prabhakaran, D., Frost, C. D. & Boothroyd, A. T. Magnetic spectrum of the two-dimensional antiferromagnet La_2CoO_4 studied by inelastic neutron scattering. *Phys. Rev. B* **82**, 184425 (2010).
27. Yao, D. X., Carlson, E. W. & Campbell, D. K. Magnetic excitations of stripes near a quantum critical point. *Phys. Rev. Lett.* **97**, 017003 (2006).
28. Kivelson, S. A. *et al.* How to detect fluctuating stripes in the high temperature superconductors. *Rev. Mod. Phys.* **75**, 1201–1241 (2003).
29. Vojta, M. Lattice symmetry breaking in cuprate superconductors: stripes, nematics, and superconductivity. *Adv. Phys.* **58**, 699–820 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Essler and E.W. Carlson for discussions. This work was supported by the Engineering and Physical Sciences Research Council of Great Britain and the Paul Scherrer Institut, Switzerland.

Author Contributions D.P. prepared and characterized the single-crystal samples. A.T.B., P.B. and P.G.F. performed the neutron scattering experiments. A.T.B. developed the theoretical model and P.B. performed the data analysis. A.T.B. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.T.B. (a.boothroyd@physics.ox.ac.uk).

Transport through modes in random media

Jing Wang¹ & Azriel Z. Genack¹

Excitations in complex media are superpositions of eigenstates that are referred to as ‘levels’ for quantum systems and ‘modes’ for classical waves. Although the Hamiltonian of a complex system may not be known or solvable, Wigner conjectured¹ that the statistics of energy level spacings would be the same as for the eigenvalues of large random matrices. This has explained key characteristics of neutron scattering spectra². Subsequently, Thouless and co-workers argued^{3,4} that the metal–insulator transition in disordered systems^{4–6} could be described by a single parameter, the ratio of the average width and spacing of electronic energy levels: when this dimensionless ratio falls below unity, conductivity is suppressed by Anderson localization⁵ of the electronic wavefunction. However, because of spectral congestion due to the overlap of modes^{7–9}, even for localized waves, a comprehensive modal description of wave propagation has not been realized. Here we show that the field speckle pattern¹⁰ of transmitted radiation—in this case, a microwave field transmitted through randomly packed alumina spheres—can be decomposed into a sum of the patterns of the individual modes of the medium and the central frequency and linewidth of each mode can be found. We find strong correlation between modal field speckle patterns, which leads to destructive interference between modes. This allows us to explain complexities of steady state and pulsed transmission of localized waves and to harmonize wave and particle descriptions of diffusion.

Modes of the field in media for which the corresponding particles freely diffuse extend throughout the sample. Energy is then readily transported to the margins of the sample, where it leaks through the boundary. Because the mode lifetime is then short and its linewidth correspondingly broad, the average spectral width of modes exceeds the average spacing between neighbouring modes^{3,4,6}: $\delta\omega > \Delta\omega$. Level repulsion of spectrally overlapping levels leads to universal fluctuations in conductance in small metallic samples¹¹ and is associated with classically chaotic motion¹² for waves inside a cavity. In contrast, the coupling of exponentially peaked modes to the environment surrounding the sample is exponentially small, so that^{3,4,6,11} $\delta\omega < \Delta\omega$. Average transmission is consequently suppressed^{3–7} while relative fluctuations of transmission¹³ or conductance¹⁴ are enhanced as a result of occasional resonant tunnelling through localized modes. The degree of spectral mode overlap $\delta \equiv \delta\omega/\Delta\omega$ often referred to as the Thouless number, is thus a fundamental localization parameter. Anderson localization has been observed for electrons¹⁵ and cold atoms¹⁶, as well as for classical waves^{14,17,18} such as light¹⁹, microwave radiation¹³, plasmons²⁰ and ultrasound^{21,22}.

A decisive step forward in the study of transport was taken with the development of the scaling theory of localization^{4,6} in which the scaling of a directly measurable quantity, the dimensional conductance at zero temperature g , was shown to depend only upon g itself and the dimensionality of the system. It was argued that the dimensionless conductance $g = G/(e^2/h)$, where G is the electronic conductance, e is the electron charge and h is Planck’s constant, is equal to the Thouless number^{4,6}, $\delta = g$. g can also be defined for classical waves via the Landauer relation, which equates the dimensionless conductance to the ensemble average of the transmittance $\langle T \rangle$. The transmittance is the sum of coefficients of total transmission T_a for the N propagating transverse incident channels with index a : $g = \langle T \rangle = \left\langle \sum_{a=1}^N T_a \right\rangle$.

The statistics of transmission in nondissipative diffusive samples are also determined by g (refs 13 and 23–26). The variance of the total transmission normalized by its ensemble average $\text{var}(s_a \equiv T_a/\langle T_a \rangle)$ and the variance of normalized intensity $\text{var}(s_{ab} = T_{ab}/\langle T_{ab} \rangle)$ are linked via the relation^{25,26} $\text{var}(s_a) = [\text{var}(s_{ab}) - 1]/2 = 2/3g$. The intensity T_{ab} corresponds to the transmission coefficient between incident and outgoing channels a and b , respectively. More generally, the full distribution of intensity and total transmission can be expressed in terms of a statistical conductance^{13,27} $g' \equiv 2/3\text{var}(s_a)$, even in strongly scattering dissipative samples. In the absence of inelastic processes g' reduces to g (refs 13, 22 and 27). Thus we expect that transmission statistics in nondissipative systems are linked to the statistical properties of the underlying modes via the relation $g' = \delta$. Confirming this relation presents a challenge because it holds only in the absence of dissipation. Therefore, measurements must either be carried out in nondissipative systems or a means must be found of obtaining these parameters from measurements in the presence of inelastic scattering, dephasing or absorption.

Measurements of the frequency variation of the speckle pattern of microwave field transmitted through random samples of alumina spheres contained in a copper tube were carried out as shown schematically in Fig. 1a. The field was detected using a short wire antenna connected to a vector network analyser. Spectra were taken over a tight grid of points on the sample output. Intensity speckle patterns at a fixed frequency such as shown in Fig. 1a were formed from these spectra. New sample configurations were realized by momentarily rotating the sample tube about its axis. From measurements over an ensemble of 200 configurations of scattering spheres for samples of length 61 cm in the frequency range 10–10.24 GHz, we obtained $g' = 0.23$. This is well below the value at the localization threshold of unity in nonabsorbing samples and corresponds to a sample of more than twice the localization length.

Energy leaks from the open ends of the sample tube and radiation is absorbed by the sample, so electromagnetic energy is not conserved and the time-evolution operator is not hermitian. Nonetheless, we expect that the time evolution of the wave following pulsed excitation can be expressed as a sum of terms, each of which is a product of a volume field speckle pattern and an exponentially decaying sinusoidal function. These terms, which are often referred to as quasi-normal modes, can form a complete set for outgoing waves²⁸. In the frequency domain, the field at the output surface can be expressed as a superposition of such modes:

$$E_j(\mathbf{r}, \omega) = \sum_n a_{n,j}(\mathbf{r}) \frac{\Gamma_n/2}{\Gamma_n/2 + i(\omega - \omega_n)} = \sum_n a_{n,j}(\mathbf{r}) \varphi_n(\omega) \quad (1)$$

Here, $a_{n,j}(\mathbf{r})$ is the spatial variation of the j th polarization component of the field for the n th mode with central frequency ω_n and linewidth Γ_n given in radians. $\varphi_n(\omega)$ is the frequency variation of the mode which is the Fourier transform of $\exp(-\frac{\Gamma_n t}{2}) \cos \omega_n t$ for $t > 0$. The average number of modes per configuration over the frequency range studied is 47. Spectra at all \mathbf{r} in a given configuration share a common set of ω_n and Γ_n , so the mode expansion in equation (1) can be fitted simultaneously to spectra at many points to give ω_n and Γ_n and the corresponding

¹Department of Physics, Queens College of the City University of New York, Flushing, New York 11367, USA.

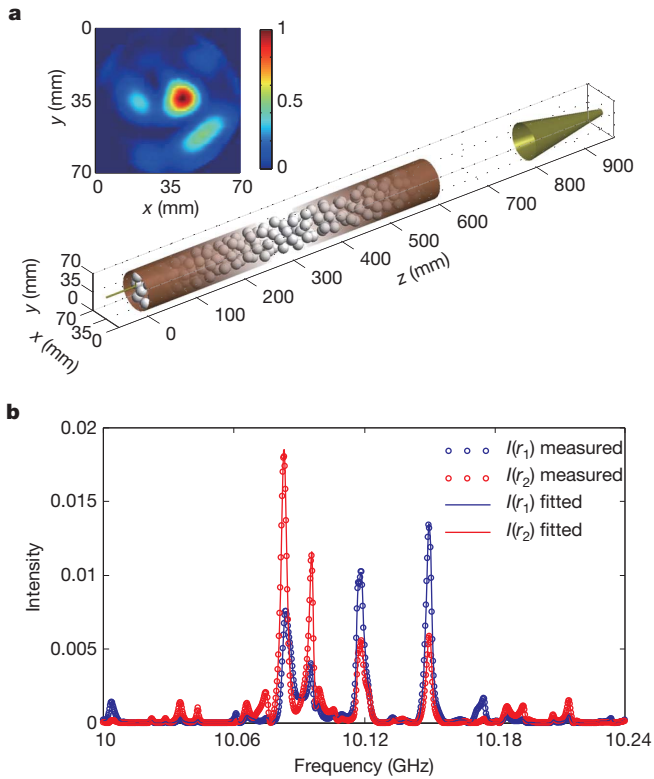


Figure 1 | Measurements of transmission through random media.

a, Microwave radiation is launched from a horn placed 20 cm before the sample. Field spectra of transmission through randomly positioned alumina spheres at low filling fraction contained in a copper tube are measured at points on a 2-mm grid over the output surface using a short wire antenna and a vector network analyser. Squaring the field at each point gives the intensity speckle pattern normalized to its peak value (inset). **b**, Intensity spectra at two detector positions r_1 and r_2 and the fit of equation (1) to the data.

mode speckle patterns. Measured intensity spectra at two points and spectra obtained from the simultaneous modal fit of spectra at 45 points are shown in Fig. 1b. The measured normalized total transmission s_a and the sum over the output surface of $|E_j(\mathbf{r}, \omega)|^2$ obtained from this global fit are shown in Fig. 2a. Excellent agreement between measurements and the global fit is obtained for spectra of both local (Fig. 1b) and integrated (Fig. 2a) transmission.

To explore the contribution of modes to total transmission, we considered a narrow frequency range around the strong peak at 10.15 GHz for the same configuration for which intensity spectra are shown in Fig. 1b. The asymmetrical shape for the line in both intensity and total transmission indicates that more than a single mode contributes to the peak. The modal analysis of the field spectra shows that three modes contribute substantially. Total transmission spectra for each of the three modes closest to 10.15 GHz are plotted in Fig. 2a. The integrated transmission for two of these modes—corresponding to the 28th and 29th modes in the spectrum starting at 10 GHz—are each greater than for the measured transmission peak, indicating that these modes interfere destructively. The intensity and phase patterns for each of these modes are shown in Fig. 2b–e. Aside from a difference in average transmission, the intensity speckle patterns of the two modes are nearly identical. At the same time, the distributions of phase shift at 10.15 GHz for the two modes are similar up to a nearly constant phase difference with average value of $\Delta\phi = 1.02\pi$. The surprising similarity between the speckle patterns for these overlapping modes suggests that these spectrally overlapping modes are formed from coupled resonances that overlap spatially within the sample^{7–9}. The similarity in the intensity speckle patterns of adjacent modes and the uniformity of the phase shift across the patterns of these modes makes interference possible between modes across the entire speckle pattern.

The modal decomposition of field spectra affords a full account of dynamic and steady-state transmission. Pulsed transmission was obtained by taking the Fourier transform of the product of transmitted field spectra $E_j(\mathbf{r}, \nu)$ and the Fourier transform of an incident Gaussian field pulse, $E_j(\mathbf{r}, \nu) \exp(-(\nu - \nu_0)^2/2\sigma^2)$. From this, we computed the

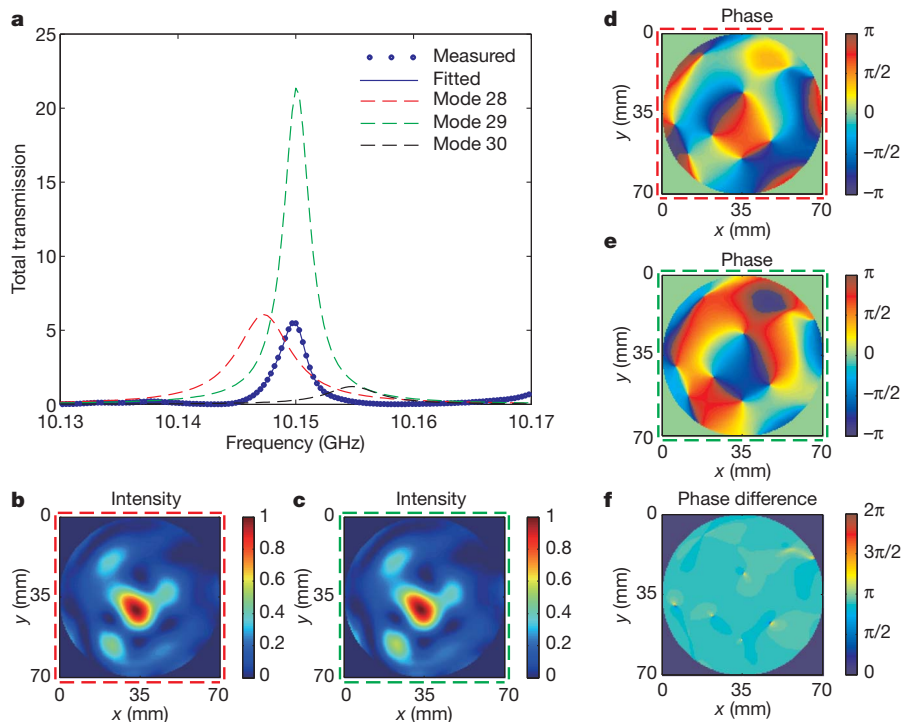


Figure 2 | Total transmission and mode speckle patterns. **a**, Three modes contribute to the asymmetric peak in the total transmission spectrum (blue) for the same configuration as in Fig. 1, mode 28 (red), mode 29 (green) and mode

30 (black). Dashed red (mode 28) and green (mode 29) lines surround their intensity speckle patterns (**b** and **c**) and phase patterns (**d** and **e**). **f**, The phase in mode 29 is shifted by nearly a constant of 1.02π rad relative to mode 28.

time–frequency spectrogram of the total transmission $T_a(t, \nu_0)$, corresponding to the variation with carrier frequency ν_0 of the sum of intensity over all points on the output surface on which measurements of field are made at delay time t from the peak of an incident Gaussian pulse: $T_a(t, \nu_0) = \sum_{\mathbf{r}} |E(\mathbf{r}, t, \nu_0)|^2$. $T_a(t, \nu_0)$ is indicated by the colour scale in the x – y plane in Fig. 3. The evolution of the spectra of total transmission is further indicated by plotting $T_a(t, \nu_0)$ normalized to the average over each spectrum for four delay times. The increasing relative weight of long-lived narrow-line modes is manifest in the decreasing number of surviving modes with substantial relative intensity. This results in an increasing variance of normalized transmission with time delay. The variation with time of the spectrogram of total transmission normalized by the peak transmission for the mode with centre frequency $\nu_n = 10.04352$ GHz can be seen in Fig. 3b to narrow and take on the Gaussian lineshape of the incident pulse. Figure 3c shows the decay of this peak. The decay rate of $1.18 \mu\text{s}^{-1}$ is essentially equal to the linewidth $\Gamma_n = 1.29$ MHz.

The average temporal variation of total transmission $\langle T_a(t) \rangle$ is obtained by integrating the time–frequency spectrogram over frequency in each configuration at different times and averaging over all configurations. The progressive suppression of transmission in time by absorption may be removed by multiplying $\langle T_a(t) \rangle$ by $\exp(t/\tau_a)$ to give $\langle T_a^0(t) \rangle = \langle T_a(t) \rangle \exp(t/\tau_a)$. The loss of energy is then due only to leakage from the sample. The decay rate $1/\tau_a = 0.0064 \text{ ns}^{-1}$ is determined from measurements of decay in a closed sample in which the

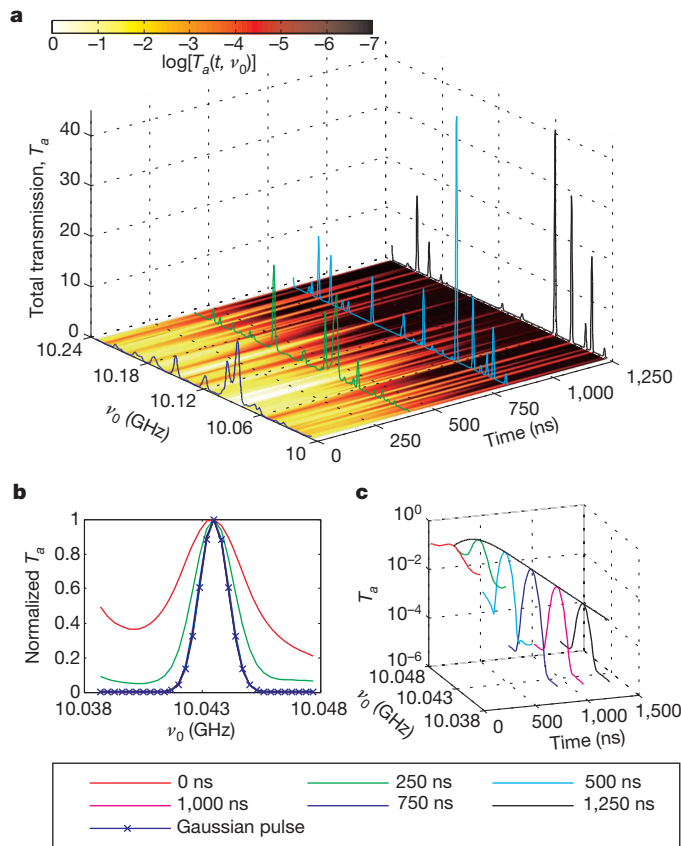


Figure 3 | Time–frequency analysis. **a**, Logarithm of time–frequency spectrogram of total transmission plotted (see colour scale) in the x – y plane. The central frequency ν_0 of the incident Gaussian pulse of linewidth $\sigma = 0.85$ MHz is scanned. Each of the four spectra of total transmission at different delay times are normalized to its average. **b**, Spectrograms of total transmission for the mode with centre frequency $\nu_n = 10.04352$ GHz are normalized by the peak transmission for different delay times. The colour of the curves in **b** and **c** indicate different delay times for the spectrograms in **b** and **c** (see legend at the bottom). All the curves overlap for delays of 500 ns and longer. **c**, The decay rate of the peak in **b** of $1.18 \mu\text{s}^{-1}$ is essentially equal to the linewidth $\Gamma_n = 1.29$ MHz.

ends of the tube are capped. The same result is obtained by transforming into the time domain using spectra computed from the modal decomposition of the field in which the impact of absorption on each mode is eliminated. The linewidth of the n th mode Γ_n is the sum of the rate at which the modal energy leaks through the sample boundaries Γ_n^0 and the absorption rate $1/\tau_a$, so the linewidth in an equivalent sample without absorption is given by: $\Gamma_n^0 = \Gamma_n - 1/\tau_a$. The complex amplitude of the mode without absorption is given by $a_{n,j}^0 = a_{n,j} \Gamma_n / \Gamma_n^0$, in agreement with results of one-dimensional simulations (this work). The decay of $\langle T_a^0(t) \rangle$, shown as the solid black curve in Fig. 4, can be seen to slow considerably with time delay^{22,29}. This reflects a broad range of modal decay rates from long-lived localized modes peaked at points remote from the sample surface^{3,4,14} to more extended short-lived modes^{7–9}.

The measured decay corrected for absorption is compared to the incoherent sum of transmission for all modes in the random ensemble $\langle \sum_n T_{a_n}^0(t) \rangle$, shown as the dashed curve in Fig. 4. $\langle \sum_n T_{a_n}^0(t) \rangle$ is substantially larger than $\langle T_a^0(t) \rangle$ at early times before it converges to $\langle T_a^0(t) \rangle$. Although transmission associated with individual modes rises with the incident pulse, transmission at early times is strongly suppressed by destructive interference of modes with strongly correlated field speckle patterns such as those shown in Fig. 2. The delayed rise of pulsed transmission seen in Fig. 4 is a remnant of particle diffusion associated with transport in weakly scattering samples and reflects the low probability of particles traversing the sample by a sequence of scattering events that are all in the forward direction. At later times, the random frequency differences between modes lead to additional random phasing between modes: $(\omega_{n+1} - \omega_n)t$. This leads to a peak in transmission after a delay time comparable to the inverse of the typical mode linewidth. By this time, destructive interference is substantially diminished and subsequently averaged pulsed transmission approaches the incoherent sum of decaying modes.

To compare δ and g' , we eliminate the effect of absorption on field spectra by substituting $a_{n,j}^0$ and Γ_n^0 for the amplitude and decay rate in the modal decomposition of equation (1). We calculate δ by identifying the average level width, or inverse Thouless time, with the average modal leakage rate $\delta\omega \equiv \tau_{\text{Th}}^{-1} \equiv (1/\Gamma_n^0)^{-1}$, and equating the average level spacing $\Delta\omega$ to the average angular frequency difference between neighbouring modes. The average over 200 sample configurations gives $\delta = \delta\omega/\Delta\omega = 0.17$. This equals the statistical conductance $g' = 0.17$, determined from the variance of intensity in spectra in which absorption is removed. δ and g' are also found to be close for an ensemble of 40 configurations of samples of length 40 cm with $\delta = 0.43$ and $g' = 0.39$. Figure 5 presents the relationship between the inverses of δ and g' , which express the degree of wave localization within the sample. The origin of the plot is the diffusive limit in which the number of modes encompassed within the linewidth diverges and fluctuations of normalized total transmission vanish^{23–26}. The red line

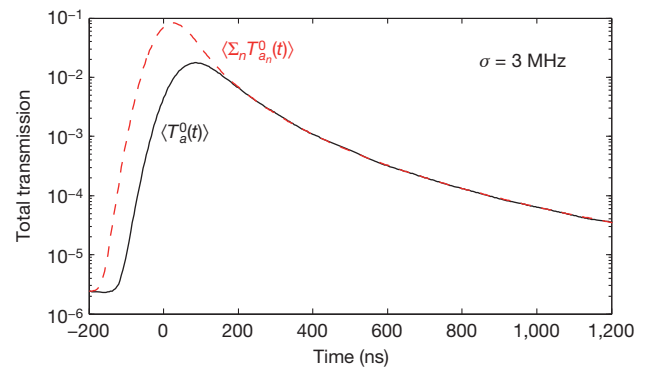


Figure 4 | Dynamics of localized waves. Semilogarithmic plot of the ensemble average of pulsed transmission and the incoherent sum of transmission due to all modes in the random ensemble. The impact of absorption is eliminated as described in the text.

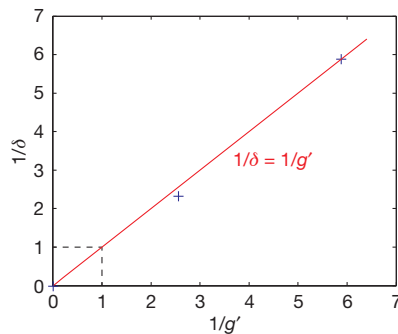


Figure 5 | Localization parameters. The degree of localization is tracked by the inverses of the mode overlap parameter, which is equivalent to the Thouless number δ and the statistical conductance g' (which reflects fluctuations in transmission). The dashed lines indicate the localization threshold at $\delta = g' = 1$. Measurements are made for sample lengths of 40 cm and 61 cm.

with slope unity is an extension of the equality of δ and g' calculated for diffusive waves^{25,26} into the localized^{13,27} regime.

We have shown that steady-state spectra and wave dynamics are explained in terms of destructive interference of strongly correlated modal speckle patterns. Notwithstanding the complex correlation found between modes, the statistics of transmission are simply related to the mode overlap parameter δ , as seen in the equality of g' and δ for diffusive and localized waves. Mode statistics are crucial for nonlinear as well as for linear phenomena and determine the threshold for narrow-line lasing in random amplifying media in systems in which linewidths and level spacings are comparable³⁰. The results presented here show that the standing waves of a random medium are the key to a full description of transport.

METHODS SUMMARY

Measurements were carried out on the microwave field transmitted through an ensemble of samples of randomly positioned alumina spheres embedded in Styrofoam spheres contained within a copper tube. The tube was 99.999% copper to minimize losses in reflection. It had a diameter of 7.2 cm and supported 30 transverse modes of the empty waveguide over the frequency range of the experiment, 10–10.24 GHz. The sample was composed of 99.95% alumina spheres of diameter 0.95 cm and refractive index 3.14 embedded in Styrofoam spheres of diameter 1.9 cm and refractive index 1.04 to achieve an alumina volume-filling fraction of 0.068. Waves were thoroughly mixed in the quasi-one-dimensional sample geometry of the experiments so that intensity statistics were essentially uniform across the output. The signal was detected by a 4-mm-long 300- μ m-diameter wire antenna translated on a 2-mm-square grid over the output surface. The separable nonlinear least-squares fitting method was used to decompose field spectra simultaneously at each of 45 positions into a sum over modes using a common set of $\omega_m \Gamma_n$. The minimum was found for the sum $\min \chi^2 = \min \sum_{p=1}^{P=45} \left| E(r_p, \nu) - \sum_{n=1}^N a_n(r_p) \phi_n \right|^2$, where r_p are the different positions at which spectra were fitted. The problem was solved to yield $\omega_m \Gamma_n$ and the mode speckle patterns by using classical nonlinear least-squares optimization methods^{31,32}, such as the Levenberg–Marquardt algorithm. Figures 3 and 4 were computed from the measured field spectra using time–frequency analysis³³ for pulses with $\sigma = 0.85$ MHz and 3 MHz, respectively, corresponding to Gaussian temporal pulse widths of $\sigma_t = 187.4$ ns and 53.1 ns.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 May 2010; accepted 12 January 2011.

1. Wigner, E. P. On a class of analytic functions from the quantum theory of collisions. *Ann. Math.* **53**, 36–67 (1951).

2. Dyson, F. J. Statistical theory of the energy levels of complex systems. I. *J. Math. Phys.* **3**, 140–156 (1962).
3. Edwards, J. T. & Thouless, D. J. Numerical studies of localization in disordered systems. *J. Phys. Chem.* **5**, 807–820 (1972).
4. Thouless, D. J. Maximum metallic resistance in thin wires. *Phys. Rev. Lett.* **39**, 1167–1169 (1977).
5. Anderson, P. W. Absence of diffusion in certain random lattices. *Phys. Rev.* **109**, 1492–1505 (1958).
6. Abrahams, E., Anderson, P. W., Licciardello, D. C. & Ramakrishnan, T. V. Scaling theory of localization: absence of quantum diffusion in two dimensions. *Phys. Rev. Lett.* **42**, 673–676 (1979).
7. Mott, N. F. & Twose, W. D. The theory of impurity conduction. *Adv. Phys.* **10**, 107–163 (1961).
8. Pendry, J. B. Quasi-extended electron states in strongly disordered systems. *J. Phys. Chem.* **20**, 733–742 (1987).
9. Bliokh, K., Yu., Bliokh, Yu. P., Freilikher, V., Genack, A. Z. & Sebbah, P. Coupling of localization mode in random media: level repulsion and necklace states. *Phys. Rev. Lett.* **101**, 133901 (2008).
10. Nye, J. F. & Berry, M. V. Dislocations in wave trains. *Proc. R. Soc. Lond. A* **336**, 165–190 (1974).
11. Altshuler, B. L. & Shklovskii, B. I. Repulsion of energy levels and conductivity of small metal samples. *Sov. Phys. JETP* **64**, 127–135 (1986).
12. Berry, M. V. Semiclassical theory of spectral rigidity. *Proc. R. Soc. Lond.* **400**, 229–251 (1985).
13. Chabanov, A. A., Stoytchev, M. & Genack, A. Z. Statistical signatures of photon localization. *Nature* **404**, 850–853 (2000).
14. Azbel, M. Ya. Eigenstates and properties of random systems in one dimension at zero temperature. *Phys. Rev. B* **28**, 4106–4125 (1983).
15. Abrahams, E. (ed.) *50 Years of Anderson Localization* (World Scientific, 2010).
16. Aspect, A. & Inguscio, M. Anderson localization of ultracold atoms. *Phys. Today* **62**, 30–35 (2009).
17. John, S. Electromagnetic absorption in a disordered medium near a photon mobility edge. *Phys. Rev. Lett.* **53**, 2169–2172 (1984).
18. Lagendijk, A. van Tiggelen, B. & Wiersma, D. S. Fifty years of Anderson localization. *Phys. Today* **62**, 24–29 (2009).
19. Schwartz, T., Bartal, G., Fishman, S. & Segev, M. Transport and Anderson localization in disordered two-dimensional photonic lattices. *Nature* **446**, 52–55 (2007).
20. Gréillon, S. *et al.* Experimental observation of localized optical excitations in random metal-dielectric films. *Phys. Rev. Lett.* **82**, 4520–4523 (1999).
21. Weaver, R. L. Anderson localization of ultrasound. *Wave Motion* **12**, 129–142 (1990).
22. Hu, H., Strybulevych, A., Page, J. H., Skipetrov, S. E. & Van Tiggelen, B. A. Localization of ultrasound in a three-dimensional elastic network. *Nature Phys.* **4**, 945–948 (2008).
23. Mello, P. A., Akkermans, E. & Shapiro, B. Macroscopic approach to correlations in the electronic transmission and reflection from disordered conductors. *Phys. Rev. Lett.* **61**, 459–462 (1988).
24. Feng, S., Kane, C., Lee, P. A. & Stone, A. D. Correlations and fluctuations of coherent wave transmission through disordered media. *Phys. Rev. Lett.* **61**, 834–837 (1988).
25. Nieuwenhuizen, Th. M. & Van Rossum, M. C. Intensity distribution of waves transmitted through a multiple scattering medium. *Phys. Rev. Lett.* **74**, 2674–2677 (1995).
26. Kogan, E. & Kaveh, M. Random-matrix-theory approach to the intensity distributions of waves propagating in a random medium. *Phys. Rev. B* **52**, R3813–R3815 (1995).
27. Stoytchev, M. & Genack, A. Z. Measurement of the probability distribution of total transmission in random waveguides. *Phys. Rev. Lett.* **79**, 309–312 (1997).
28. Ching, E. S. C., Leung, P. T., Suen, W. M., Tong, S. S. & Young, K. Waves in open systems: eigenfunction expansions. *Rev. Mod. Phys.* **70**, 1545–1554 (1998).
29. Zhang, Z. Q., Chabanov, A. A., Cheung, S. K., Wong, C. H. & Genack, A. Z. Dynamics of localized waves: pulsed microwave transmissions in quasi-one-dimensional media. *Phys. Rev. B* **79**, 144203 (2009).
30. Cao, H. *et al.* Random laser action in semiconductor powder. *Phys. Rev. Lett.* **82**, 2278–2281 (1999).

Acknowledgements This research was supported by the National Science Foundation (DMR0907285).

Author Contributions J.W. improved the apparatus, took the data, developed the modal decomposition and the time–frequency analysis algorithms, analysed the data and contributed to writing the paper. A.Z.G. largely conceived and directed the research and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.Z.G. (genack@qc.edu).

METHODS

Experimental method. Spectra of the microwave field transmission coefficient through random waveguides were taken using a vector network analyser. The sample is a copper tube filled with randomly positioned 99.95% alumina spheres of diameter 0.95 cm and refractive index 3.14 embedded in Styrofoam spheres of diameter 1.9 cm and refractive index 1.04 to achieve an alumina filling fraction of 0.068. The tube was 99.999% copper to minimize losses in reflection. The tube had a diameter of 7.2 cm and supported 30 transverse propagating modes over the frequency range of the experiment of 10 to 10.24 GHz. The sample was locally three-dimensional within the reflecting tube with length much greater than its diameter. The random waveguide was thus similar to a microscopically disordered wire in electronics. The wave incident at different parts of the input surface of this quasi-one-dimensional sample was thoroughly mixed on the output surface and intensity statistics were essentially uniform across the output except for fluctuations arising from the finite number of transverse modes. Radiation was strongly scattered within the sample over the frequency range of the experiment which was just above the first Mie resonance of the alumina spheres²⁷. The value of δ varied by less than 10% over this frequency range in the sample with $L = 61$ cm.

The field output of the network analyser was amplified before it was directed at the sample through a horn placed on the axis of the sample tube as shown in Fig. 1. The transmitted field was detected by a wire antenna translated on a 2-mm-square grid over the output surface. The antenna was 4 mm long and 300 μ m in diameter and was connected to the network analyser through a cable. The signal detected was the integral over the volume of the antenna of the component of the field directed along the wire axis. The dimensions of the wire were much smaller than the microwave wavelength in air of about 3 cm. Speckle patterns at each frequency were obtained from field spectra taken at each point on the output surface. The speckle patterns shown in Figs 1 and 2 were obtained using the two-dimensional Whittaker–Shannon sampling theorem. Ensembles of sample realizations were created by rotating the sample tube for several seconds after each measurement.

Modes of an open system. The field of a mode of an open system may be factorized into space- and time-varying terms after pulse excitation. The time-varying term for the n th mode is $\cos \omega_n t \exp(-\frac{\Gamma_n}{2}t)$ for $t > 0$ and zero for $t < 0$. The decay rate of energy for the n th mode is Γ_n . The Fourier transform of the single sided exponential decay is the factor $\varphi_n(\omega) = \frac{\Gamma_n/2}{\Gamma_n/2 + i(\omega - \omega_n)}$. Given that the form of $\varphi_n(\omega)$ is independent of dimension, we expect that the influence of absorption upon the mode linewidth and strength will similarly not depend on dimension. We therefore explored the impact of absorption in a simpler one-dimensional geometry using transfer matrix simulations. We found that the product $a_{n,j}\Gamma_n$ is independent of absorption and that spectra found without absorption are replicated to high accuracy by substituting $a_{n,j}(\mathbf{r})\Gamma_n/\Gamma_n^0$ for $a_{n,j}(\mathbf{r})$ and $\Gamma_n^0 = \Gamma_n - 1/\tau_a$ for Γ_n in equation (1). This substitution was used to find the spectra that would be obtained in the equivalent nonabsorbing samples. The absorption rate $1/\tau_a$ was determined from measurements of the energy decay rate in a sample with reflecting end caps in which leakage from the sample was suppressed. Spatial modes of the medium are expected to be orthogonal over the volume of the sample²⁸ but not across the output plane. This makes it possible for different modes to have similar speckle patterns.

Separable nonlinear least-squares fitting. We used the separable nonlinear least-squares fitting method to decompose field spectra simultaneously at each of 45 positions into a sum over modes using a common set of ω_n, Γ_n . The minimum is

$$\text{found for the sum } \min \chi^2 = \min \sum_{p=1}^{P=45} \left| E(r_p, v) - \sum_{n=1}^N a_n(r_p) \varphi_n \right|^2, \text{ where } r_p \text{ are the}$$

different positions at which the spectra were fitted. The problem was solved to yield the ω_n, Γ_n and the mode speckle patterns by using classical nonlinear least-squares optimization methods, such as the Levenberg–Marquardt algorithm. In matrix format, this is:

$$\min \chi^2 = \min \|E - \Phi a\|^2 \quad (2)$$

where each column of the matrix E represents a spectrum $\text{Re}E(v_1), \text{Im}E(v_1), \dots, \text{Re}E(v_M), \text{Im}E(v_M)$ at a single position r_p and each column of a represents the corresponding complex amplitudes of the modes, $\text{Re}a_1, \text{Im}a_1, \dots, \text{Re}a_N, \text{Im}a_N$ at position r_p . The elements of matrix Φ are $\Phi_{2m-1, 2n-1} = \text{Re}\varphi_n(v_m)$, $\Phi_{2m-1, 2n} = -\text{Im}\varphi_n(v_m)$, $\Phi_{2m, 2n-1} = \text{Im}\varphi_n(v_m)$ and $\Phi_{2m, 2n} = \text{Re}\varphi_n(v_m)$ for the n th mode at frequency point m . The matrix a can very nearly be expressed as $a = \Phi^\dagger E$, where Φ^\dagger is the Moore–Penrose pseudo-inverse of the matrix Φ (refs 31 and 32). Eliminating the linear coefficient a , the fitting problem becomes:

$$\min \chi^2 = \min \|(I - \Phi \Phi^\dagger)E\|^2 \quad (3)$$

where $\|\dots\|$ indicates the Euclidean norm of the matrix and I is the unit matrix. The problem can be solved using classical nonlinear least-squares optimization methods, such as the Levenberg–Marquardt algorithm. Fitting was performed simultaneously for $P = 45$ points. The set of ω_n and Γ_n obtained from this global fit were then used to solve for the spatial mode amplitudes $a_{n,j}(r_p)$ for all measured points.

Time–frequency analysis. To reveal the impact of modes on the dynamics of waves transmitted through random media, we computed the temporal evolution of the field as a function of the carrier frequency v_0 of an incident Gaussian pulse. The spectrum of the incident pulse $E_0(\mathbf{r}, t, v_0) \approx \exp(-(t^2/2\sigma_t^2)\cos(2\pi v_0 t))$ is a Gaussian in the frequency domain $E_0(t) \approx \exp(-(v - v_0)^2/2\sigma^2)$, where $\sigma = (2\pi\sigma_t)^{-1}$. Figures 3 and 4 were computed from the measured field spectra using time–frequency analysis³³ for pulses with $\sigma = 0.85$ MHz and 3 MHz, respectively, corresponding to Gaussian temporal pulse widths of $\sigma_t = 187.4$ ns and 53.1 ns. The increasing impact of long-lived narrow-line modes can be seen in the time–frequency spectra in Fig. 3. The narrow linewidth of the incident pulse makes it possible to determine the central frequencies of long-lived modes with high accuracy. These are used as initial values of parameters in the fit of the mode expansion in equation (1) to the measured spectrum.

31. Golub, G. H. & Pereyre, V. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Probl.* **19**, R1–R26 (2003).
32. Sima, D. M. & Huffel, S. V. Separable nonlinear least squares fitting with linear bound constraints and its application in magnetic resonance spectroscopy data quantification. *J. Comput. Appl. Math.* **203**, 264–278 (2007).
33. Cohen, L. *Time-Frequency Analysis* (Prentice Hall PTR, 1995).

Eocene global warming events driven by ventilation of oceanic dissolved organic carbon

Philip F. Sexton^{1,2,†}, Richard D. Norris¹, Paul A. Wilson², Heiko Pälike², Thomas Westerhold³, Ursula Röhl³, Clara T. Bolton^{2,†} & Samantha Gibbs²

'Hyperthermals' are intervals of rapid, pronounced global warming known from six episodes within the Palaeocene and Eocene epochs (~65–34 million years (Myr) ago)^{1–13}. The most extreme hyperthermal was the ~170 thousand year (kyr) interval² of 5–7 °C global warming³ during the Palaeocene–Eocene Thermal Maximum (PETM, 56 Myr ago). The PETM is widely attributed to massive release of greenhouse gases from buried sedimentary carbon reservoirs^{1,3,6,11,14–17}, and other, comparatively modest, hyperthermals have also been linked to the release of sedimentary carbon^{3,6,11,16,17}. Here we show, using new 2.4-Myr-long Eocene deep ocean records, that the comparatively modest hyperthermals are much more numerous than previously documented, paced by the eccentricity of Earth's orbit and have shorter durations (~40 kyr) and more rapid recovery phases than the PETM. These findings point to the operation of fundamentally different forcing and feedback mechanisms than for the PETM, involving redistribution of carbon among Earth's readily exchangeable surface reservoirs rather than carbon exhumation from, and subsequent burial back into, the sedimentary reservoir. Specifically, we interpret our records to indicate repeated, large-scale releases of dissolved organic carbon (at least 1,600 gigatonnes) from the ocean by ventilation (strengthened oxidation) of the ocean interior. The rapid recovery of the carbon cycle following each Eocene hyperthermal strongly suggests that carbon was re-sequestered by the ocean, rather than the much slower process of silicate rock weathering proposed for the PETM^{1,3}. Our findings suggest that these pronounced climate warming events were driven not by repeated releases of carbon from buried sedimentary sources^{3,6,11,16,17}, but, rather, by patterns of surficial carbon redistribution familiar from younger intervals of Earth history.

Earth's climate attained its warmest state of the past 80 Myr during the early Palaeogene period (~45–65 Myr ago) (Fig. 1a). Centred within this interval at 56 Myr ago, the PETM marks a rapid excursion of 2.5‰ to >3.0‰ in the carbon isotope composition ($\delta^{13}\text{C}$) of marine carbonates^{1,3} and terrestrial organic matter concurrent with widespread dissolution of carbonate sediments^{1,3} and deep ocean warming by 6–7 °C (ref. 3). The source of carbon triggering this large perturbation to the exchangeable carbon reservoirs at Earth's surface (ocean, atmosphere and biosphere) is widely acknowledged to have come from an 'external' sedimentary reservoir(s)^{1,3,6,11,14–17}. A long-standing hypothesis attributes this event to the release of massive amounts of methane from sedimentary gas hydrate deposits^{1,6,11,14–16}, with attendant oxidation to CO_2 causing climate warming through a greenhouse feedback³. Because of the relative rapidity of its onset, and the estimated quantity of sedimentary-sourced carbon involved, the PETM has attracted interest as a natural analogue to the current anthropogenic perturbation to Earth's surficial carbon cycle.

In recent years, six transient warming events comparable in character to the PETM, but less extreme in magnitude and duration, have been discovered throughout the early Palaeogene (at about 65.2, 58.2, 53.7,

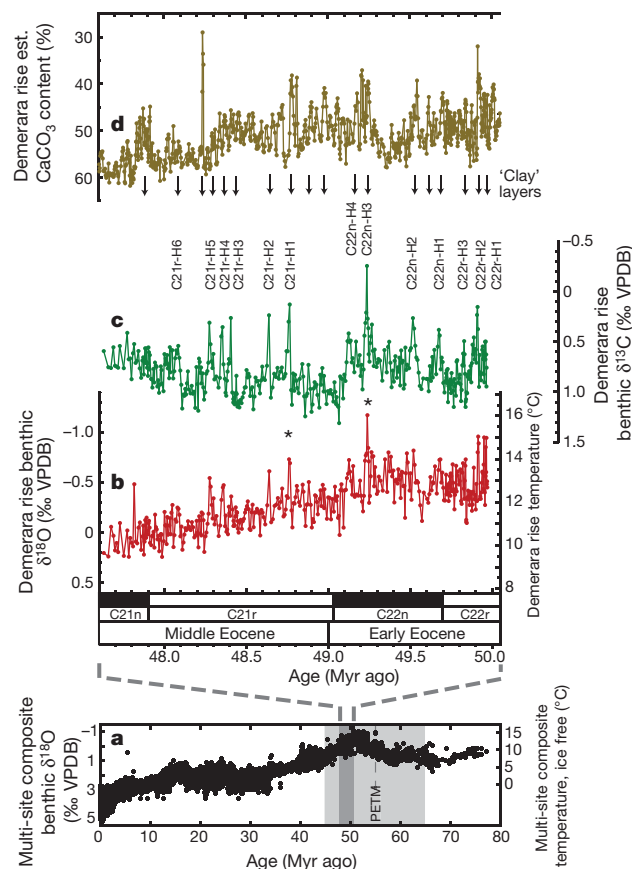


Figure 1 | High-resolution records across the early Eocene to middle Eocene transition from ODP Site 1258, Demerara rise, tropical Atlantic. **a**, Deep sea benthic foraminifer composite $\delta^{18}\text{O}$ data for the past 80 Myr from multiple sites^{3,8,10} (full list of data sources in Supplementary Note 1). PETM, Palaeocene–Eocene Thermal Maximum. Light grey shaded area, early Palaeogene period (45–65 Myr ago). Dark grey shaded box, study interval for this work. $\delta^{18}\text{O}$ -temperature scale is for ice-free conditions³ (–1.27‰ VPDB). **b–d**, Data from ODP Site 1258, Demerara rise, with an astronomically calibrated age model (Methods). **b**, Benthic foraminifer $\delta^{18}\text{O}$ and calculated palaeotemperatures (Methods). All benthic $\delta^{18}\text{O}$ data in **a** and **b** are corrected to equilibrium calcite. **c**, Benthic foraminifer $\delta^{13}\text{C}$. Note the repeated, rapid excursions to lower values. Events numbered (according to the magnetochron within which they occur) are those with a $\delta^{13}\text{C}$ excursion exceeding 0.7‰. Black arrows mark the stratigraphic positions of clay-rich CaCO_3 dissolution horizons ($n = 18$) made visible by their brown colour in the Demerara rise sedimentary sequence, 13 of which correspond precisely to the numbered, rapid excursions to lower benthic $\delta^{13}\text{C}$ values in **c**. The remaining 5 clay layers are also associated with trends to lower $\delta^{13}\text{C}$, but of noticeably smaller magnitude. **d**, Estimated CaCO_3 content of sediment. Asterisks denote individual 'hyperthermal' events shown in more detail in Fig. 2.

¹Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA. ²National Oceanography Centre Southampton, University of Southampton, European Way, Southampton SO14 3ZH, UK. ³MARUM — Center for Marine Environmental Sciences, University of Bremen, Leobener Strasse, 28359 Bremen, Germany. [†]Present addresses: Department of Earth and Environmental Sciences, The Open University, Milton Keynes MK7 6AA, UK (P.F.S.); Departamento de Geología, Universidad de Oviedo, 33005 Oviedo, Spain (C.T.B.).

53.2, 52.5 and 41.8 Myr ago^{4–13}). These ‘hyperthermals’⁴ have not been found during the later and cooler part of the Eocene (after ~40 Myr ago) (Fig. 1a). The number of these hyperthermals, and the relative temporal spacing between some of them, is suggestive of the operation of common mechanisms^{5,6,11}. Although various mechanisms have been proposed for the origin of these comparatively modest events^{5,7,12}, it is widely considered that they too were triggered by large-scale releases of carbon from sedimentary reservoirs^{3,6,11,16,17}, probably methane hydrates^{3,6,11,16}. Yet, the typical magnitude of the $\delta^{13}\text{C}$ anomalies across known hyperthermals (up to ~1.0‰) is not unusual for $\delta^{13}\text{C}$ records from the better chronicle Neogene period and, more importantly, these anomalies are not so large that they necessitate a source of carbon that is as extremely isotopically depleted as methane hydrate ($\delta^{13}\text{C} = -60\text{‰}$).

A new 2.4-Myr-long benthic foraminifer stable isotope record (Supplementary Discussion) from Demerara rise in the tropical western Atlantic Ocean (Ocean Drilling Program Site 1258) shows that events exhibiting the characteristics of previously identified hyperthermals are abundant in the Palaeogene deep-sea palaeoceanographic record. Spanning the early Eocene to middle Eocene transition (47.6–50.0 Myr ago), these records indicate that average temperatures in the early Eocene equatorial Atlantic were 12–14 °C at about 3,000 m (ref. 8) palaeo-water depth (Fig. 1b). This warm deep ocean state is punctuated by 13 excursions to lower $\delta^{18}\text{O}$ values indicative of short-term warming by ~2–4 °C (Figs 1b, 2; events numbered in Fig. 1c), each lasting about 40 kyr and spaced about 100–400 kyr apart (Supplementary Discussion). These transient warming events develop very rapidly (<5–10 kyr) and decay to ‘background’ temperatures more slowly (over ~30 kyr) (Fig. 2).

Our accompanying benthic foraminifer $\delta^{13}\text{C}$ record shows that each warming event is marked by an excursion to lower $\delta^{13}\text{C}$ values by 0.7–1.0‰ (Fig. 1c). Where planktic records are available for the most prominent negative excursions in the benthic $\delta^{13}\text{C}$ record, these also reveal negative $\delta^{13}\text{C}$ excursions of 0.6–0.9‰ (Fig. 2), suggesting that these events represent whole-ocean decreases in $\delta^{13}\text{C}$ of total dissolved inorganic carbon.

All of the benthic $\delta^{13}\text{C}$ excursions are associated with brown clay-rich layers at Demerara rise (arrows in Fig. 1), a hallmark of other reported early Palaeogene hyperthermals^{1,6,7,10,12,13}, that suggests increased calcium carbonate (CaCO_3) dissolution. Clay layers occur in other parts of the record as well, associated with $\delta^{13}\text{C}$ excursions of smaller magnitude than our numbered excursions (Fig. 1c). This suggests that the largest hyperthermals (those numbered in Fig. 1c) represent prominent end-members in a spectrum of carbon cycle perturbations that scale up from orbitally paced ‘background’ cycles. Records of estimated CaCO_3 concentrations from Demerara rise (palaeo-water depth ~3 km; ref. 8)

and sites in the southern Atlantic and central Pacific oceans (palaeo-water depths ~3 and 2 km, respectively; Supplementary Discussion) reveal synchronous, globally widespread decreases in CaCO_3 accumulation (Supplementary Fig. 2e–g). Calculations of CaCO_3 dissolution associated with the clay layers in the Atlantic and Pacific (Supplementary Discussion) indicate that CaCO_3 dissolution events (Fig. 3b–d) are coeval with the large negative benthic $\delta^{13}\text{C}$ excursions at Demerara rise (Fig. 3a). This correspondence strongly suggests that global increases in deep ocean acidity were associated with intervals of isotopically ‘light’ carbon pervading the world’s oceans.

These perturbations to oceanic carbon cycling could have been fuelled by carbon input from a number of different sources. Besides methane hydrates, erosion of sedimentary organic carbon (C_{org}) could supply carbon with a $\delta^{13}\text{C}$ composition (about –25‰) that is less depleted than that of methane (–60‰). A third potential source is redistribution of carbon within the exchangeable reservoirs at Earth’s surface (ocean, atmosphere and biosphere). Redistribution of isotopically light carbon among the internal, readily exchangeable reservoirs should produce $\delta^{13}\text{C}$ excursions with fundamentally different size, shape, duration and recovery trends than carbon input from an external, sedimentary reservoir.

Mass balance considerations (assuming masses of Eocene carbon reservoirs similar to modern) dictate that the ~1‰ $\delta^{13}\text{C}$ excursions we observe are compatible with injection of a mass of carbon from methane hydrates (–60‰) of about 650 gigatonnes (Gt). But this mass of carbon is far too small to have driven our corresponding deep ocean warmings of 2–4 °C (Figs 1b, 2), particularly given that deep ocean warming of 6–7 °C at the PETM³ was triggered by release of ~3,000 to 6,000 Gt of carbon¹⁷. The large and widespread increases in CaCO_3 dissolution associated with our hyperthermals also appear to demand a much larger input of carbon than the 650 Gt permitted by a methane hydrate source.

The $\delta^{13}\text{C}$ composition of sedimentary C_{org} (about –25‰) permits a much larger carbon release (~1,600 Gt) to be accommodated within the observed 1‰ $\delta^{13}\text{C}$ excursions of hyperthermals, but problems also exist with this explanation. The ~140-kyr residence time of carbon in Earth’s exchangeable reservoirs^{14,15} suggests that the liberation of a substantial quantity of isotopically light carbon from an external, deeply buried source should have a long imprint on oceanic and atmospheric $\delta^{13}\text{C}$ (approaching the residence time)¹⁸ until the excess carbon is ultimately sequestered back into the long-term sedimentary carbon reservoir by weathering reactions¹⁸. This behaviour is demonstrated at the PETM, which shows a $\delta^{13}\text{C}$ excursion ~170 kyr in duration² with a very long recovery ‘tail’^{1,3,5,11} that produces a very asymmetrical excursion and reflects the long timescales required for weathering to draw down this external CO_2 and return it to the sedimentary reservoir as CaCO_3 .

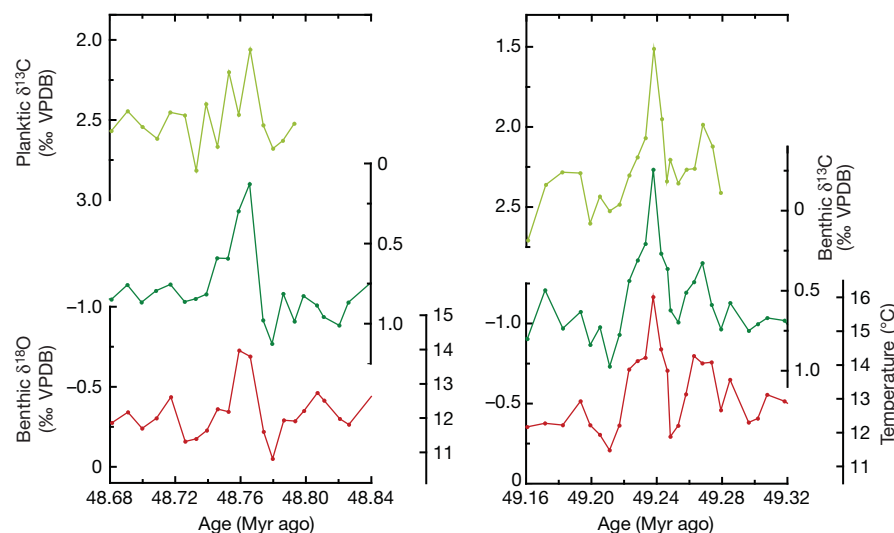


Figure 2 | Details of stable isotope data from ODP Site 1258 across two ‘hyperthermal’ events. Isotope data are from Fig. 1; the two events are marked by asterisks in Fig. 1. Labels on the x-axes are at 40-kyr spacing, highlighting the similarity of the duration (that is, ‘trough to trough’) of these events to the obliquity frequency. We define the ‘onset’ phase of events as ‘trough to peak’, with their ‘recovery’ being ‘peak to trough’. Red, benthic foraminifer $\delta^{18}\text{O}$. Dark green, benthic $\delta^{13}\text{C}$. Light green, planktic $\delta^{13}\text{C}$.

However, our $\delta^{13}\text{C}$ excursions are too short (only ~ 40 -kyr total duration), and return to pre-excursion values too fast (within 30 kyr; Figs 1c, 2) and to a post-excursion baseline that is isotopically no different from pre-excursion, to be compatible with large-scale liberation of isotopically light carbon from a buried reservoir. The recoveries seen for other hyperthermals^{5,6,8,10–13} are also too rapid to represent sequestration of substantial amounts of sedimentary-sourced carbon.

The incompatibility of hyperthermals with external, sedimentary carbon sources indicates that the carbon that fuelled them was probably redistributed within the readily exchangeable reservoirs at Earth's surface. Although some of this carbon may have come from the terrestrial biosphere, possibly acting as a feedback to a primary source, we are unable to test this possibility with our data at present. However, the oceans are by far the largest of these exchangeable reservoirs; the present-day oceans contain 13 times as much carbon as the combined biosphere and atmosphere—all told about 40,000 Gt of carbon. Today, this mass of oceanic carbon comprises about 37,500 Gt of dissolved inorganic carbon ($\delta^{13}\text{C} \approx 0\text{‰}$), 30 Gt of particulate organic carbon ($\delta^{13}\text{C} \approx -25\text{‰}$) and 700–1,800 Gt of dissolved organic carbon (DOC; $\delta^{13}\text{C} \approx -25\text{‰}$). Redistribution of a mass of isotopically light carbon equivalent to the entire modern DOC reservoir ($\sim 1,600$ Gt C) would be required to produce the $\sim 1.0\text{‰}$ $\delta^{13}\text{C}$ excursions seen for our events and previously reported ones^{5,6,8,10–13}, but the DOC reservoir size would probably grow, perhaps many fold, under more oxygen-depleted deep ocean conditions. Within the oceans, the two principal reservoirs for storing large quantities of carbon are marginal basins or the abyssal ocean. Because of its sheer size, the deep abyssal reservoir of carbon has long been implicated in the regular fluctuations (~ 100 parts per million by volume, p.p.m.v.) of atmospheric CO_2 and $\sim 5^\circ\text{C}$ surface temperature changes seen across the late Pleistocene glacial cycles. It has recently become clear that the abyssal Southern Ocean has played a pivotal role in Pleistocene carbon cycling through its gradual amassing of respired CO_2 during glacials^{19–21} and subsequent release via increased ventilation^{19–21} during deglaciation.

Several observations suggest that the source of CO_2 fuelling Eocene hyperthermals was the abyssal ocean. First, our dissolution records provide clues to the location of CO_2 storage in the exchangeable carbon reservoirs. Numerical modelling experiments indicate that CaCO_3 dissolution should be most intense close to the source of carbon release¹⁵. In our estimates of CaCO_3 dissolution, dissolution intensity appears to be consistently highest in the southern Atlantic (Fig. 3e) compared to other sites. This finding raises the possibility that the abyssal reservoir of carbon was located in the Southern Ocean. Second, all of our hyperthermals have a duration of about 40 kyr (Fig. 1b, c, Fig. 2). The similarity of this period to the 41-kyr obliquity cycle suggests that the forcing for individual hyperthermal events had its origin at high latitudes. This observation is consistent with an obliquity pacing of high latitude surface ocean stratification controlling carbon ventilation (via oxygenation), as proposed for the last deglaciation²².

If the source of CO_2 driving Eocene hyperthermals was the deep ocean, a number of observations suggest that our Eocene sites may be somewhat analogous to intermediate-depth sites during Pleistocene glacials that were located above a deeper CO_2 storage reservoir^{19,23}. First, the absence before our hyperthermals of either gradually decreasing benthic $\delta^{13}\text{C}$ or increasing dissolution intensity (Figs 1c, 2) indicates that our mid-depth (2–3 km) sites did not 'sense' the gradual ($\sim 10^4$ – 10^5 year) build-up of a CO_2 reservoir until its final release. Second, early Eocene sediments from lower abyssal depths reveal lithological indications for relatively long intervals of carbon storage followed by short, intermittent intervals of carbon release (Supplementary Discussion).

The relatively long ~ 100 -kyr or 400-kyr intervals between CO_2 release events (Fig. 1) would suggest an apparent means by which a large Eocene abyssal reservoir of DOC could intermittently grow. Yet increases in the DOC reservoir are difficult to achieve solely by long residence times of deep water, owing to progressive bacterial

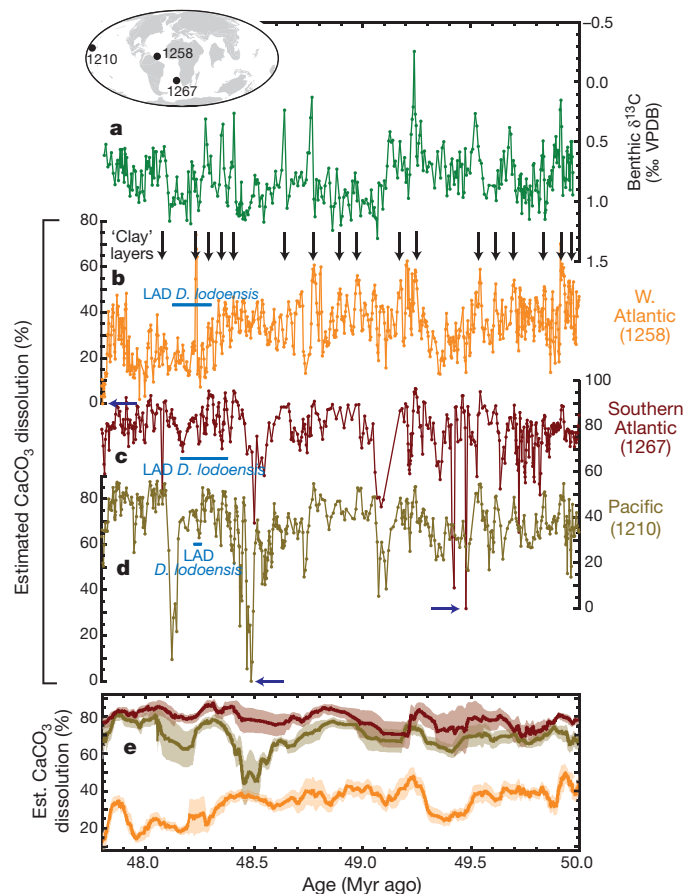


Figure 3 | Eocene records of benthic foraminifer $\delta^{13}\text{C}$ and CaCO_3 dissolution. **a**, Benthic $\delta^{13}\text{C}$ from Fig. 1c (Demerara rise, ODP Site 1258, tropical Atlantic). **b–d**, Records of calculated CaCO_3 dissolution from ODP Site 1258 (**b**), ODP Site 1267, southern Atlantic (**c**), ODP Site 1210, Pacific (**d**), and smoothed versions (25-point running mean) of all three (**e**). Shaded envelopes in **e** are 1σ variance about the mean. Simultaneous increases in CaCO_3 dissolution occur at all three sites and are marked by visible, brown 'clay' layers in sediment cores at each of the sites (marked by solid black arrows). Blue arrows mark our assumed 'zero-dissolution points' for each site (that is, the primary, unaltered fractional component of CaCO_3 in the sediment; Supplementary Fig. 2 and Supplementary Discussion). Light blue horizontal bars mark the interval at each site within which the last appearance datum (LAD) of the nannofossil *Discoaster lodoensis* occurs, providing a primary point of stratigraphic correlation between sites. Inset map from ODSN Plate Tectonic Reconstruction Service.

respiration primarily of young, labile DOC components during the water mass ageing process²⁴. However, under anoxic conditions, DOC may build up. For example, concentrations of DOC in the anoxic depths ($>2,000$ m) of the modern Black Sea are ~ 2.5 times those of the global deep ocean²⁵. An early Eocene decoupling of marine burial rates of C_{org} versus pyrite indicates that the ocean during this interval regularly experienced pronounced anoxia²⁶. We suggest that this persistent anoxia, perhaps in the severely undersampled early Eocene abyss, and at least partly driven by the lower solubility of oxygen at warmer abyssal temperatures (Fig. 1a, b), would have promoted the build-up of a reservoir of DOC large enough to have fuelled hyperthermal carbon releases of $\sim 1,600$ Gt. Anoxia in the Eocene abyss is consistent with modelling projections for the next few millennia indicating a large reduction in abyssal ventilation (by up to 75%)²⁷ in response to climate warming induced by an 'anthropogenic' atmospheric partial pressure of CO_2 (1,700 p.p.m.v.) similar to that of the early Eocene.

Further support for the existence of a larger-than-modern Eocene deep (>3 km) ocean reservoir of DOC is found in surface-to-deep profiles of seawater temperature and $\delta^{13}\text{C}$ reconstructed from multiple species of planktic and benthic foraminifera (Supplementary Fig. 4). In

comparison to recent core-top data, Eocene data reveal a greater total surface-to-seafloor $\delta^{13}\text{C}$ difference along with a steeper vertical gradient of $\delta^{13}\text{C}$ of dissolved inorganic carbon versus temperature (Supplementary Fig. 4). These two observations are respectively consistent with a deeper mean C_{org} remineralization depth and an intensification of organic carbon production in the upper ocean. The resultant greater Eocene vertical export flux of carbon may have allowed the unremineralized DOC component to build up to higher-than-modern levels. Furthermore, the much greater species diversity of calcareous nannofossils relative to diatoms throughout the early Palaeogene²⁸ may have accelerated the abyssal transfer of C_{org} from the surface owing to the greater ballasting efficiency of CaCO_3 relative to opal²⁹. A major switch to greater species diversity of diatoms relative to calcareous nannofossils (at ~ 40 Myr ago)²⁸ coincides with the last documented early Cenozoic 'hyperthermal' event¹⁰.

Although the PETM must have been fuelled by carbon injection from an external, sedimentary source(s), we suggest that astronomically paced changes in ocean ventilation (oxygenation) of DOC may explain the numerous less extreme hyperthermals that are being discovered^{14–13} throughout the early Palaeogene. CO_2 ventilation is a viable mechanism with which to sustain the repeated, frequent carbon releases that define hyperthermals over a geologically extended interval (Fig. 1b, c)^{14–13} because the recharge times of the abyssal ocean (via decomposition of organic matter) are rapid compared to the slow, multi-million-year filling of sedimentary methane hydrate reservoirs³⁰. The extraordinarily warm mean state of the Palaeogene oceans (Fig. 1a, b) also raises the question as to whether large-scale methane hydrate reservoirs would even have existed at this time³⁰. Despite the extreme nature of these repeated Palaeogene climate warming events, our findings indicate that they were driven by changes in storage of carbon within the oceans, in a manner somewhat familiar from younger intervals of Earth history.

METHODS SUMMARY

$\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ data were generated using mono-specific analyses of the benthic foraminifers *Cibicides subpiratus* and *Cibicides eocanus* at the National Oceanography Centre, University of Southampton, Southampton. High resolution Eocene records of CaCO_3 concentrations were calculated, by regression, using sediment physical property data and measurements of sediment CaCO_3 content. We developed an astronomically calibrated age model for Site 1258 (Demerara rise). Detailed chronologies for sites 1267 and 1210 were generated by tuning clear, correlatable events in their respective CaCO_3 (%) records to the astronomically calibrated CaCO_3 (%) record from Site 1258.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 August 2010; accepted 12 January 2011.

1. Zachos, J. C. *et al.* Rapid acidification of the ocean during the Paleocene-Eocene thermal maximum. *Science* **308**, 1611–1615 (2005).
2. Röhl, U., Westerhold, T., Bralower, T. J. & Zachos, J. C. On the duration of the Paleocene-Eocene thermal maximum (PETM). *Geochim. Geophys. Geosyst.* **8**, Q12002 (2007).
3. Zachos, J. C., Dickens, G. R. & Zeebe, R. E. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451**, 279–283 (2008).
4. Thomas, E., Zachos, J. C. & Bralower, T. J. in *Warm Climates in Earth History* (eds Huber, B., MacLeod, K. & Wing, S.) 132–160 (Cambridge Univ. Press, 2000).
5. Cramer, B. S., Wright, D. D., Kent, D. V. & Aubry, M. P. Orbital climate forcing of $\delta^{13}\text{C}$ excursions in the late Paleocene-early Eocene (chons C24n-C25n). *Paleoceanography* **18**, 1097, doi:10.1029/2003PA000909 (2003).
6. Lourens, L. J. *et al.* Astronomical pacing of late Palaeocene to early Eocene global warming events. *Nature* **435**, 1083–1087 (2005).
7. Petrizzo, M. R. An early late Paleocene event on Shatsky Rise, northwest Pacific Ocean (ODP Leg 198): evidence from planktonic foraminiferal assemblages. *Proc. ODP Sci. Res.* **198**, 1–29 (2005).
8. Sexton, P. F., Wilson, P. A. & Norris, R. D. Testing the Cenozoic multisite composite $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ curves: new monospecific Eocene records from a single locality, Demerara Rise (Ocean Drilling Program Leg 207). *Paleoceanography* **21**, PA2019, doi:10.1029/2005PA001253 (2006).
9. Westerhold, T. *et al.* On the duration of magnetochrons C24r and C25n and the timing of early Eocene global warming events: implications from the Ocean Drilling Program Leg 208 Walvis Ridge depth transect. *Paleoceanography* **22**, PA2201, doi:10.1029/2006PA001322 (2007).

10. Edgar, K. M., Wilson, P. A., Sexton, P. F. & Sugauma, Y. No extreme bipolar glaciation during the main Eocene calcite compensation shift. *Nature* **448**, 908–911 (2007).
11. Nicolo, M. J., Dickens, G. R., Hollis, C. J. & Zachos, J. C. Multiple early Eocene hyperthermals: their sedimentary expression on the New Zealand continental margin and in the deep sea. *Geology* **35**, 699–702 (2007).
12. Quillevér, F., Norris, R. D., Kroon, D. & Wilson, P. A. Transient ocean warming and shifts in carbon reservoirs during the early Danian. *Earth Planet. Sci. Lett.* **265**, 600–615 (2008).
13. Stap, L. *et al.* High-resolution deep-sea carbon and oxygen isotope records of Eocene Thermal Maximum 2 and H2. *Geology* **38**, 607–610 (2010).
14. Dickens, G. R., O'Neil, J. R., Rea, D. K. & Owen, R. M. Dissociation of oceanic methane hydrate as a cause of the carbon isotope excursion at the end of the Paleocene. *Paleoceanography* **10**, 965–971 (1995).
15. Dickens, G. R. Methane oxidation during the late Palaeocene thermal maximum. *Bull. Soc. Geol. Fr.* **171**, 37–49 (2000).
16. Dickens, G. R. Rethinking the global carbon cycle with a large, dynamic and microbially mediated gas hydrate capacitor. *Earth Planet. Sci. Lett.* **213**, 169–183 (2003).
17. Panchuk, K., Ridgwell, A. & Kump, L. R. Sedimentary response to Paleocene-Eocene Thermal Maximum carbon release: a model-data comparison. *Geology* **36**, 315–318 (2008).
18. Archer, D. Fate of fossil fuel CO_2 in geologic time. *J. Geophys. Res.* **110**, C09S05, doi:10.1029/2004JC002625 (2005).
19. Hodell, D. A., Venz, K. A., Charles, C. D. & Ninnemann, U. S. Pleistocene vertical carbon isotope and carbonate gradients in the South Atlantic sector of the Southern Ocean. *Geochim. Geophys. Geosyst.* **4** (1), 1004, doi:10.1029/2002GC000367 (2003).
20. Toggweiler, J. R., Russell, J. L. & Carson, S. R. Midlatitude westerlies, atmospheric CO_2 , and climate change during the ice ages. *Paleoceanography* **21**, PA2005, doi:10.1029/2005PA001154 (2006).
21. Skinner, L. C., Fallon, S., Waelbroeck, C., Michel, E. & Barker, S. Ventilation of the deep Southern Ocean and deglacial CO_2 rise. *Science* **328**, 1147–1151 (2010).
22. Sigman, D. M., de Boer, A. M. & Haug, G. H. in *Past and Future Changes of the Oceanic Meridional Overturning Circulation: Mechanisms and Impacts* (eds Schmittner, A., J., Chiang, H. C. & Hemming, S. R.) 335–350 (AGU Geophysical Monograph 173, American Geophysical Union, 2007).
23. Marchitto, T. M., Lehman, S. J., Ortiz, J. D., Fluckiger, J. & van Geen, A. Marine radiocarbon evidence for the mechanism of deglacial atmospheric CO_2 rise. *Science* **316**, 1456–1459 (2007).
24. Hansell, D. A. & Carlson, C. A. Deep-ocean gradients in the concentration of dissolved organic carbon. *Nature* **395**, 263–266 (1998).
25. Ducklow, H. W., Hansell, D. A. & Morgan, J. A. Dissolved organic carbon and nitrogen in the western Black Sea. *Mar. Chem.* **105**, 140–150 (2007).
26. Kurtz, A. C., Kump, L. R., Arthur, M. A., Zachos, J. C. & Paytan, A. Early Cenozoic decoupling of the global carbon and sulfur cycles. *Paleoceanography* **18**, 1090, doi:10.1029/2003PA000908 (2003).
27. Schmittner, A., Oeschies, A., Matthews, H. D. & Galbraith, E. D. Future changes in climate, ocean circulation, ecosystems, and biogeochemical cycling simulated for a business-as-usual CO_2 emission scenario until year 4000 AD. *Glob. Biogeochem. Cycles* **22**, GB1013, doi:10.1029/2007GB002953 (2008).
28. Falkowski, P. G. *et al.* The evolution of modern eukaryotic phytoplankton. *Science* **305**, 354–360 (2004).
29. Klaas, C. & Archer, D. E. Association of sinking organic matter with various types of mineral ballast in the deep sea: implications for the rain ratio. *Glob. Biogeochem. Cycles* **16** (4), 1116, doi:10.1029/2001GB001765 (2002).
30. Buffett, B. & Archer, D. Global inventory of methane hydrate: sensitivity to changes in the deep ocean. *Earth Planet. Sci. Lett.* **227**, 185–199 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Bolshaw for laboratory assistance and the shipboard party and crew of Ocean Drilling Program (ODP) Leg 207 for a successful drilling expedition. We thank H. Brinkhuis, G. Dickens, G. Foster, M. Huber, S. Kirtland, D. Kroon, L. Kump, E. Rohling and J. Zachos for discussions. This research used samples and data provided by the ODP. ODP (now IODP) is sponsored by the US NSF and participating countries under the management of JOI, Inc. We thank W. Hale and A. Wülbels (IODP) for assistance with sediment core sampling. Financial support for this research was provided by a European Commission Marie Curie Outgoing International Fellowship (P.F.S.), a Leverhulme Trust Fellowship (P.F.S.), a Natural Environment Research Council UK ODP grant (P.A.W. and P.F.S.), a Philip Leverhulme Prize (H.P.), the DFG-Leibniz Center for Surface Process and Climate Studies at the University of Potsdam, and the DFG (U.R. and T.W.).

Author Contributions P.F.S. and P.A.W. designed and instigated the research. P.F.S. and C.T.B. picked foraminifera. P.F.S. and P.A.W. generated stable isotope records. P.F.S. and H.P. generated the estimated CaCO_3 content records and constructed age models. P.F.S. conducted stratigraphic correlations between the various drill sites. T.W. and U.R. modified the spliced sedimentary section at Demerara rise. S.G. generated biostratigraphic data for Demerara rise. P.F.S. and R.D.N. wrote the manuscript. P.A.W., H.P., T.W. and U.R. commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.F.S. (P.F.Sexton@open.ac.uk).

METHODS

Stable isotope data. Benthic foraminifera were picked from the 250–400 µm size fraction. Stable isotope data were generated at the National Oceanography Centre, University of Southampton, Southampton, using a Europa Geo 20–20 mass spectrometer equipped with an automatic carbonate preparation system. For each analysis, between 4 and 11 mono-specific specimens were analysed after ultrasonic cleaning in deionized water. All data are reported relative to the Vienna Pee Dee Belemnite standard (VPDB). External analytical precision, based on replicate analyses of in-house standards calibrated to NBS-19, is better than $\pm 0.1\%$ for $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$.

Early Eocene through to middle Eocene sediments at ODP Site 1258 (Demerara rise) host well preserved benthic foraminifera⁸. *Cibicidoides subspiratus* was used as the primary species for stable isotope analysis. However, in certain intervals, notably in the younger part of our record (after 48.5 Myr ago), a decline in abundance of *C. subspiratus* dictated that we also use *C. eoceanus*. Although modern core top samples show significant inter-species stable isotope offsets between certain species of *Cibicidoides*³¹, our calculated inter-species offsets for the two Eocene species of *Cibicidoides* (from paired mono-specific analyses, $n = 168$) are statistically indistinguishable from zero for both $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ ($\delta^{18}\text{O} = -0.031$ ($1\sigma = 0.128$); $\delta^{13}\text{C} = 0.009$ ($1\sigma = 0.133$)). This provides justification for our strategy of using these two species in combination. Benthic foraminiferal taxonomy follows that of refs 32 and 33.

All $\delta^{18}\text{O}$ data are corrected to equilibrium calcite using the *Oridorsalis* spp. – *Cibicidoides* spp. correction factor of $+0.28\%$ derived from an assessment of isotopic offsets for Eocene benthic foraminifera³⁴. Use here of this correction factor is based on evidence (summarized in ref. 8) that Eocene *Oridorsalis umbonatus* secreted its calcite in $\delta^{18}\text{O}$ equilibrium with sea water. Palaeotemperatures were calculated using equation (1) of ref. 35. This equation provides excellent agreement with a global core top *Cibicidoides* spp. $\delta^{18}\text{O}$ calibration for a temperature range from 0 to 7 °C (ref. 35). In keeping with the view that there were no significant continental ice sheets during the early to early middle Eocene^{36,37}, our $\delta^{18}\text{O}$ records can be interpreted in terms of change in bottom water temperature at Demerara rise.

Age models. Drilling at Site 1258 recovered an unusually continuous, relatively expanded (sedimentation rates up to 2.5 cm kyr⁻¹) sedimentary section across the lower to middle Eocene boundary with unprecedented magneto-, bio- and cyclostratigraphic control for this interval^{38–40}. We developed an astronomically calibrated age model for Site 1258 by tuning clear and pronounced obliquity (~40-kyr) cycles in the magnetostratigraphically calibrated magnetic susceptibility (MS) time series (Supplementary Fig. 1a) to computed obliquity cycles from the most recent astronomical solution⁴¹. Our tuning strategy is validated by the presence in our tuned MS time series of substantial power (and coherency with our tuning target) at frequencies of 23, 21 and 19 kyr (precession), 54 kyr (obliquity) and 96 kyr (eccentricity) (Supplementary Fig. 1b), even though these frequencies were not used in the tuning process. An absolute calibration to the astronomical solution is, at present, unattainable for the early to middle Eocene because of limitations in the precision of computed orbital variations pre-40 Myr ago⁴¹ and because of significant uncertainties in radiometric age calibrations^{42,43}. Consequently, we develop a 'floating' astronomically tuned timescale, using the C21r/C22n magnetochron boundary (age = 49.037 Myr ago⁴²) identified at Site 1258³⁹ as our anchor point.

Age models for the estimated CaCO₃ (%) records for ODP 1267 (Walvis ridge) and ODP 1210 (Shatsky rise) are derived by correlating the stratigraphic position of a biostratigraphic datum (LAD of the nannofossil *Discoaster lodoensis*) at both

of these sites to its position at Site 1258 (Supplementary Table 1). Further age model refinement was achieved by tuning clear, correlatable events in CaCO₃ (%) records from Sites 1267 and 1210 to the astronomically calibrated CaCO₃ (%) record from Site 1258 (Supplementary Tables 2 and 3). This tuning strategy was held as relaxed as possible, using a minimum of tie points.

Eocene records of CaCO₃ concentrations. We use sediment colour and magnetic susceptibility data in conjunction with measurements of sediment CaCO₃ content³⁸ to calculate, by regression, sediment CaCO₃ content at high resolution at Demerara rise (Fig. 1d; Supplementary Fig. 2e). We use comparable data^{44,45} to perform the same regressions for two other deep sea drill sites (ODP 1267, Walvis ridge, southern Atlantic and ODP 1210, Shatsky rise, central Pacific) (Supplementary Fig. 2f and g). At all three sites, estimated CaCO₃ concentrations prove to be a good predictor of measured CaCO₃ concentrations (Supplementary Fig. 2a–c). For methods for calculating CaCO₃ dissolution from our Eocene CaCO₃ records, see Supplementary Information.

- Curry, W. B., Slowey, N. C. & Lohmann, G. P. Oxygen and carbon isotopic fractionation of aragonitic and calcitic benthic foraminifera on Little Bahama Bank, Bahamas. *Eos* **74**, 368 (1993).
- Tjalsma, R. C. & Lohmann, G. P. *Paleocene-Eocene Bathyal and Abyssal Benthic Foraminifera from the Atlantic Ocean* (Micropaleontol. Spec. Publ. Ser., Vol. 4, Micropaleontol. Proj., New York, 1983).
- van Morkhoven, F. P. C. M., Berggren, W. A. & Edwards, A. S. Cenozoic cosmopolitan deep water benthic foraminifera. *Bull. Cent. Rech. Explor. Prod. Elf-Aquitaine* **11** (Pau, France, 1986).
- Katz, M. E. et al. Early Cenozoic benthic foraminiferal isotopes: species reliability and interspecies correction factors. *Paleoceanography* **18**, 1024, doi:10.1029/2002PA000798 (2003).
- Bemis, B. E., Spero, H. J., Bijma, J. & Lea, D. W. Reevaluation of the oxygen isotopic composition of planktonic foraminifera: experimental results and revised paleotemperature equations. *Paleoceanography* **13**, 150–160 (1998).
- Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693 (2001).
- Browning, J. V., Miller, K. G. & Pak, D. K. Global implications of lower to middle Eocene sequence boundaries on the New Jersey coastal plain: the icehouse cometh. *Geology* **24**, 639–642 (1996).
- Shipboard Scientific Party, 2004. Site 1258. *Proc. ODP Init. Rep.* 207, 1–117 doi:10.2973/odp.proc.ir.207.105.2004 (2004).
- Suganuma, Y. & Ogg, J. G. Campanian through Eocene magnetostratigraphy of Sites 1257–1261, ODP Leg 207, Demerara Rise (western equatorial Atlantic). *Proc. ODP Sci. Res.* 207 (2006); available at (http://www-odp.tamu.edu/publications/207_SR/102/102.htm).
- Westerhold, T. & Röhl, U. High resolution cyclostratigraphy of the early Eocene — new insights into the origin of the Cenozoic cooling trend. *Clim. Past* **5**, 309–327 (2009).
- Laskar, J. et al. A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.* **428**, 261–285 (2004).
- Berggren, W., Kent, D. & Swisher, C. III. in *Geochronology Time Scales and Global Stratigraphic Correlation* (ed. Berggren, W.) 129–212 (Society for Sedimentary Geology, Tulsa, 1995).
- Machlus, M., Hemming, S. R., Olsen, P. E. & Christie-Blick, N. Eocene calibration of geomagnetic polarity time scale reevaluated: evidence from the Green River Formation of Wyoming. *Geology* **32**, 137–140 (2004).
- Shipboard Scientific Party, 2004. Site 1267. *Proc. ODP Init. Rep.* 208, 1–77 doi:10.2973/odp.proc.ir.208.108.2004 (2004).
- Shipboard Scientific Party, 2002. Site 1210. *Proc. ODP Init. Rep.* 198, 1–89 doi:10.2973/odp.proc.ir.198.106.2002 (2002).

The role of crustal quartz in controlling Cordilleran deformation

Anthony R. Lowry¹ & Marta Pérez-Gussinyé²

Large-scale deformation of continents remains poorly understood more than 40 years after the plate tectonic revolution¹. Rock flow strength and mass density variations both contribute to stress, so both are certain to be important, but these depend (somewhat nebulously) on rock type, temperature and whether or not unbound water is present². Hence, it is unclear precisely how Earth material properties translate to continental deformation zones ranging from tens to thousands of kilometres in width, why deforming zones are sometimes interspersed with non-deforming blocks and why large earthquakes occasionally rupture in otherwise stable continental interiors. An important clue comes from observations that mountain belts and rift zones cyclically form at the same locations despite separation across vast gulfs of time³ (dubbed the Wilson tectonic cycle), accompanied by inversion of extensional basins⁴ and reactivation of faults and other structures formed in previous deformation events⁵. Here we show that the abundance of crustal quartz, the weakest mineral in continental rocks², may strongly condition continental temperature and deformation. We use EarthScope seismic receiver functions⁶, gravity and surface heat flow measurements⁷ to estimate thickness and seismic velocity ratio, v_p/v_s , of continental crust in the western United States. The ratio v_p/v_s is relatively insensitive to temperature but very sensitive to quartz abundance^{8,9}. Our results demonstrate a surprising correlation of low crustal v_p/v_s with both higher lithospheric temperature and deformation of the Cordillera, the mountainous region of the western United States. The most plausible explanation for the relationship to temperature is a robust dynamical feedback, in which ductile strain first localizes in relatively weak, quartz-rich crust, and then initiates processes that promote advective warming, hydration and further weakening. The feedback mechanism proposed here would not only explain stationarity and spatial distributions of deformation, but also lend insight into the timing and distribution of thermal uplift¹⁰ and observations of deep-derived fluids in springs¹¹.

Separating thermal and compositional influences on lithospheric rheology, understanding the roles of crust and mantle in lithospheric stability, and explaining structural reactivation and the Wilson tectonic cycle are among the primary goals of EarthScope¹, a major research equipment initiative to illuminate solid Earth processes using dense seismic and geodetic arrays. EarthScope's transportable seismic array has collected data at more than 1,000 sites and will eventually sample the entire continental United States at ~70-km spacing.

Compressional (v_p) and shear (v_s) seismic velocity fields are somewhat ambiguous tools for separating compositional variations from thermal effects, because they are sensitive to both. However, the ratio v_p/v_s is very sensitive to compositional variations in crustal rocks, and particularly the silica content (Fig. 1a, after ref. 8). Weighted regression of density versus v_p/v_s for continental rock types tracks the line from granite to gabbro, the silicic and mafic end-members of continental crystalline rocks. Variation in v_p/v_s for a large range of temperature⁹ is comparatively small. Inspection of v_p/v_s for constituent minerals (Fig. 1b) reveals that rock compositional dependence predominantly

reflects sensitivity to quartz, because mafic minerals exhibit v_p/v_s that is similar to (or slightly lower than) feldspar and mica, the other major constituents of granite.

A geophysical tool for mapping quartz concentration in crustal rocks would illuminate crustal deformation processes, because flow strength of crustal rocks is also sensitive to quartz. Mineral flow strength² for the lesser of dislocation and diffusion creep is shown in Fig. 1c; the strength of mineral constituents bounds that of rock aggregates (and weaker constituents are favoured). Hence the key factors influencing flow viscosity are temperature, water fugacity and abundance of quartz.

The ratio v_p/v_s is not widely used to estimate crustal composition because accurate estimation is difficult. Body wave tomograms, for example, often have very different path sampling for shear and compressional waves. Receiver-function studies of differences in P- and mode-converted S-wave arrival times from the Moho are sensitive to both crustal thickness (H) and v_p/v_s (K), so both can be estimated from amplitude stacks of seismic arrivals at times predicted for a range of possible H and K (ref. 12). However, v_p/v_s is poorly determined in receiver-function analyses when crustal depths are aliased, so it is commonly treated as a nuisance parameter if seismometer spacing exceeds a few tens of kilometres. Indeed, EarthScope Automated Receiver Survey (EARS⁶) estimates of v_p/v_s behave very much like random noise (Fig. 2a and c). An example H - K parameter stack from Nevada (Fig. 2d) illustrates some of the problems: stack amplitude maxima are elongate in the K dimension, and real-Earth violation of the method's implicit assumptions (a one-dimensional medium with a single impedance contrast) results in smearing of stacked phase arrival energy and multiple maxima.

To improve resolution of bulk crustal v_p/v_s , we modified EARS H - K parameter stacks using two likelihood filters: one derived from spatial statistics (using optimal interpolation), and another from gravity modelling of the crustal thickness and v_p/v_s fields (see Methods). Crustal thickness and v_p/v_s parameters trade off unfavourably in both the receiver-function stack (Fig. 2d) and the gravity modelling (Fig. 2e). Combining the information content from the two data sets is advantageous, because the parameter confidence axes are approximately orthogonal, resulting in a 'cross-hairs' on the model parameter space (and hence a much improved estimate of v_p/v_s , shown in Fig. 3).

Variations in v_p/v_s have been attributed to both compositional variations and pore fluid pressure effects. High (that is, greater than hydrostatic) pore fluid pressure can significantly raise v_p/v_s , but it would also lower density. In the western United States, composition dominates, as evidenced by the positive sign of the density parameter relating v_p/v_s to observed gravity (grey line in Fig. 1a; see Methods). Figure 3 exhibits high v_p/v_s in several well-studied mafic provinces, including the Snake River plain (where the Yellowstone hotspot has intruded massive quantities of basalt¹³) and oceanic-derived terranes in the Cascadia forearc¹⁴ and the Great Valley forearc basin¹⁵. Crust beneath the Idaho and Coastal Range granitic batholiths has anomalously low v_p/v_s , as it does under eastern parts of the Sierra Nevada and Baja batholiths and the Cascade volcanoes. The $^{87}\text{Sr}/^{86}\text{Sr} = 0.706$

¹Department of Geology, Utah State University, Logan, Utah 84322-4505, USA. ²Department of Earth Sciences, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.

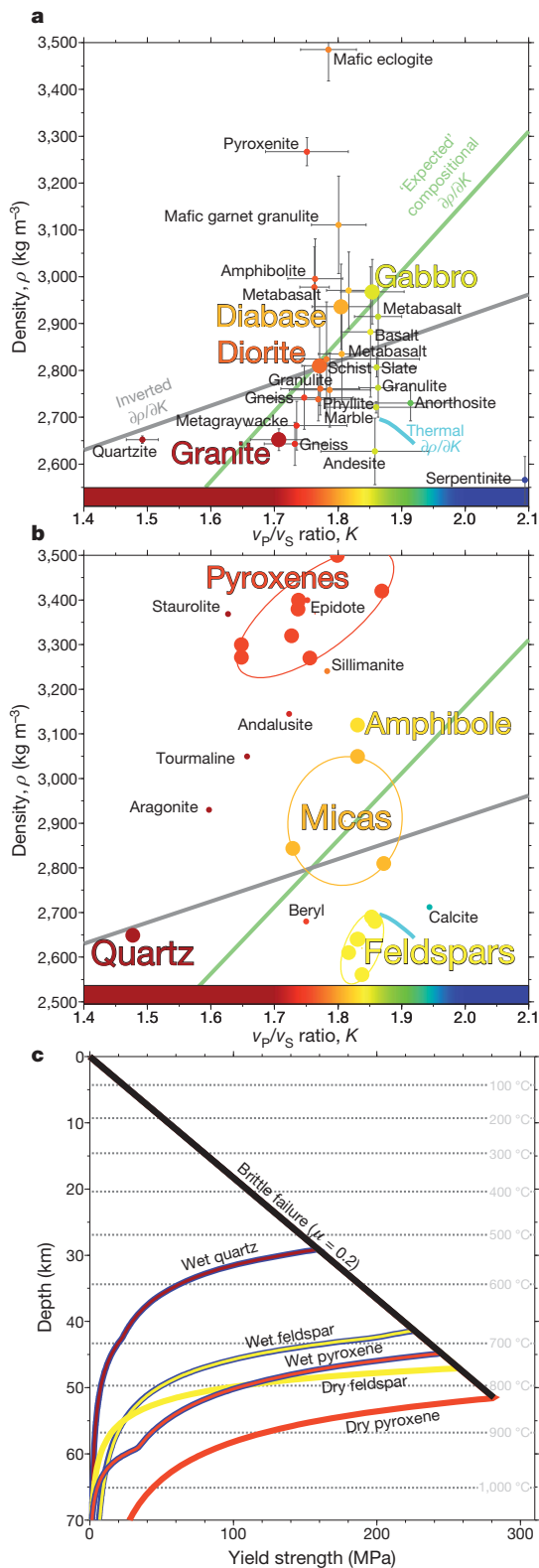


Figure 1 | Laboratory measurements of rock properties. **a**, Density ρ versus velocity ratio v_p/v_s for rocks⁸; bars are 1σ . Green line is a weighted regression of laboratory measurements; grey line depicts slope inverted from geophysical data (see Methods). Cyan curve shows feldspar temperature dependence for a 20–900 °C range⁹; other minerals behave similarly. **b**, Density versus v_p/v_s for minerals⁸ shows quartz dominates the compositional relationship. **c**, Flow strength for crustal mineral constituents², assuming 10^{-14} s⁻¹ strain rate, 1 mm grain size, and a geotherm from the Colorado plateau. Brittle-field failure assumes a frictional coefficient $\mu = 0.2$. Temperature, quartz abundance and water fugacity determine whether the lower crust flows.

isopleth^{16,17}, a contour of isotopic measurements thought to mark the boundary between young oceanic-derived crust and older, Precambrian continental crust, tracks a sharp boundary between high v_p/v_s to the west and low to the east where the isopleth is well-sampled (California and Idaho).

The variation of v_p/v_s exhibits a remarkable correspondence to the Cenozoic tectonic strain history. The Basin and Range and Rio Grande rifts have relatively low v_p/v_s ratio, particularly along the margins where modern strain rates are highest. The Rocky Mountain provinces (where Laramide thick-skin tectonism was substantial but modern strain-rates are low) also have low v_p/v_s . Deformation-resistant Cordilleran blocks such as the Great Valley, Columbia plateau and Colorado plateau have higher v_p/v_s , as does the stable continental interior, especially in the northern Great Plains.

Perhaps most interesting is the strong inverse correlation observed between v_p/v_s and surface heat flow⁷ (compare Fig. 3 with Fig. 4a). Quantitatively, the correlation coefficient r is -0.56 , significant at more than 26σ confidence, for gravity fields derived from thermal and v_p/v_s models (see Methods and Supplementary Fig. 2). This is surprising, in that the direct effect of temperature on seismic velocities (for example, cyan line in Fig. 1a) would entail high v_p/v_s in warm regions. Thermal conductivity is sensitive to composition¹⁸, but also to temperature, such that granite has 20% greater conductivity than gabbro at the surface, equal conductivity at ~ 225 °C and 30% lower conductivity at 700 °C. Hence refractory conduction through the crust, in which three-dimensional heat transfer is favoured in lateral zones of higher conductivity, should yield an opposite-signed relation to that which we observe, and the shallower variations will partially offset the opposite-signed deeper refraction yielding a net-small effect on surface heat flow. Radiogenic heat production is also sensitive to composition, and the relationship predicted between v_p/v_s and heat flow variation has the correct sign, but also would be small relative to measured heat flow variations. Radial v_s anisotropy¹⁹ probably perturbs the signal, and has been observed to be strongest in rifted corridors where advective heat transfer is also greatest. However, the pattern of anisotropy variation is very different to that of v_p/v_s , and the 0–5% measured range of crustal anisotropy roughly maps to the ~ 0.08 1σ uncertainty of the v_p/v_s estimates.

Inverse correlation of v_p/v_s to heat flow makes sense if we consider that advective processes dominate Cordilleran heat flow variations²⁰. Advective processes include lithospheric thinning via rifting²¹, asthenospheric return flow over mantle drips²², and associated processes of magmatism and fluid flux, which are amplified by warming-induced release of water bound in lower-crustal minerals²³. Higher temperatures and water both serve to decrease ductile strength, so puzzling observations like the Wilson cycle, basin inversion and structural reactivation may arise from a geodynamical feedback engendered by variations in crustal silica: ductile strain initially localizes where quartz concentrations are greatest and the crust is weakest; strain induces lithospheric thinning (in extension) or Rayleigh-Taylor instability (in contraction) which promotes advective heat transfer; the resulting warming hydrates the crust and further weakens the lithosphere, feeding back on the strain localization.

Spatial distributions of deep-derived fluid flux, inferred from surface sampling of $^3\text{He}/^4\text{He}$ and $\delta^{13}\text{C}$ isotopic ratios^{1,11}, lend support to this hypothesis: deep-derived fluid flux is found almost exclusively in regions of low crustal v_p/v_s , opposite to the relationship expected if pore fluid pressure were a significant fraction of the signal⁶. Observations of high CO_2 fluxes¹¹ are consistent with the expectation that CO_2 and water must be present in combination to buffer against resorption of these highly reactive volatiles into crystalline rock²³. These fluids may be an important vector for Cordilleran advective heat transfer, and hence it is conceivable that the relationship of crustal quartz abundance to deformation proposed here would not be possible without the water cycled into the lithosphere during subduction cycles. Of course, not all advective processes begin with surface strain: the Yellowstone-Snake River plain

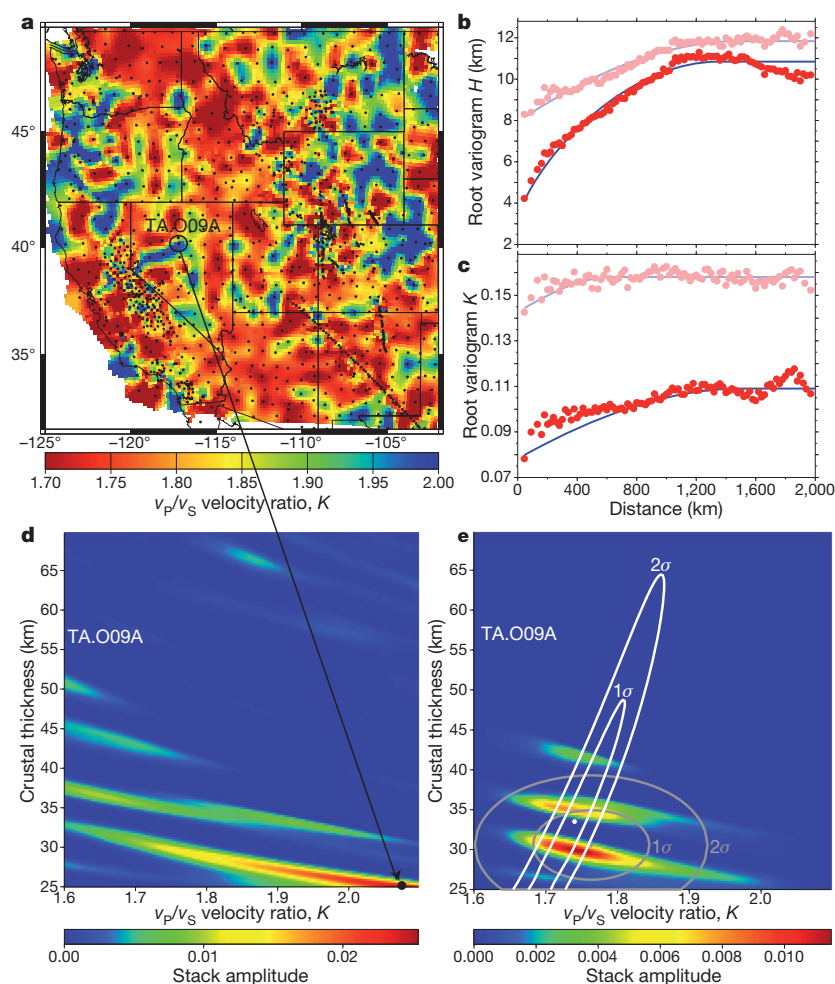


Figure 2 | Likelihood filtering of crustal thickness and v_p/v_s . **a**, Raw EARS⁶ estimates of v_p/v_s . **b**, Binned (red circles) and modelled (blue line) root-varigrams (root mean squared-difference as a function of distance) of crustal thickness measurements; light colours are raw EARS estimates; full colour saturation, our final model. **c**, Root-varigrams of v_p/v_s . **d**, EARS⁶ H - K stack for Transportable Array site O09A ($H = 25$ km; $K = 2.08$). **e**, O09A stack after likelihood filtering ($H = 29.75$; $K = 1.76$). 1-2 σ contours of optimal interpolation (grey) and gravity (white) models used to generate likelihood filters are also shown.

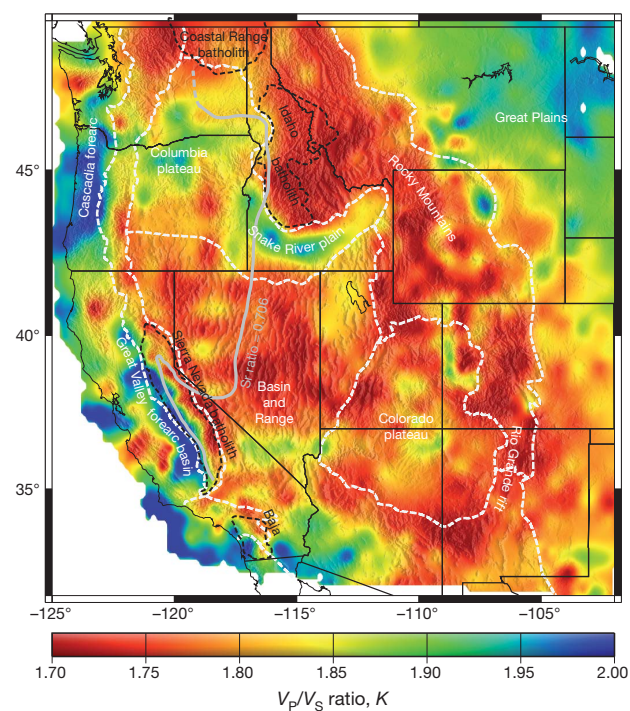


Figure 3 | Bulk crustal v_p/v_s of the western United States. Dashed white lines are physiographic province boundaries; dashed black outlines large granitic batholiths; solid grey line is the approximate $^{87}\text{Sr}/^{86}\text{Sr} = 0.706$ isotopic isopleth.

system responded to deep-mantle rather than lithosphere dynamics. But high correlation of lithospheric temperature to crustal composition suggests the feedback mechanism proposed here can be important, and advective heating promoted by deep fluid flux may also lend new insight into the distribution and timing of Cordilleran thermal uplift¹⁰.

Crustal thickness (Fig. 4b) is also inversely correlated with both heat flow ($r = -0.32$) and v_p/v_s ($r = -0.14$). These correlations are highly significant (at $>12\sigma$ and 5σ confidence, respectively), and expected if our proposed dynamical feedback is correct, but coefficients are substantially lower than for v_p/v_s and temperature. A likely reason for weaker correlation of crustal thickness is suggested by the relationship of Cordilleran strain history to lithospheric effective elastic thickness, T_e (Fig. 4c), estimated from an optimized isostatic analysis of gravity and topography²⁴. Modern crustal thickness reflects both orogenic thickening before 30 Myr ago and thinning by subsequent rifting. Values of $T_e > 30$ km imply a contribution of mantle strength²⁵, and whereas extended provinces westward of Sevier-belt thin-skinned contraction exhibit no modern mantle strength, the uppermost mantle under Laramide thick-skinned thrusts remains strong today. Extension has been relatively weak in the Laramide foreland where the mantle is strong. As Laramide thickening focused in regions of more abundant crustal quartz, the net result is thicker crust associated with lower v_p/v_s (whereas in regions of low mantle strength, the opposite is true). There is evidence of latest Quaternary normal faulting throughout the Laramide foreland provinces, however, leaving open the possibility of future weakening. Regardless, the potential for v_p/v_s mapping to help resolve ambiguities in crustal versus mantle contributions to total lithospheric strength is intriguing.

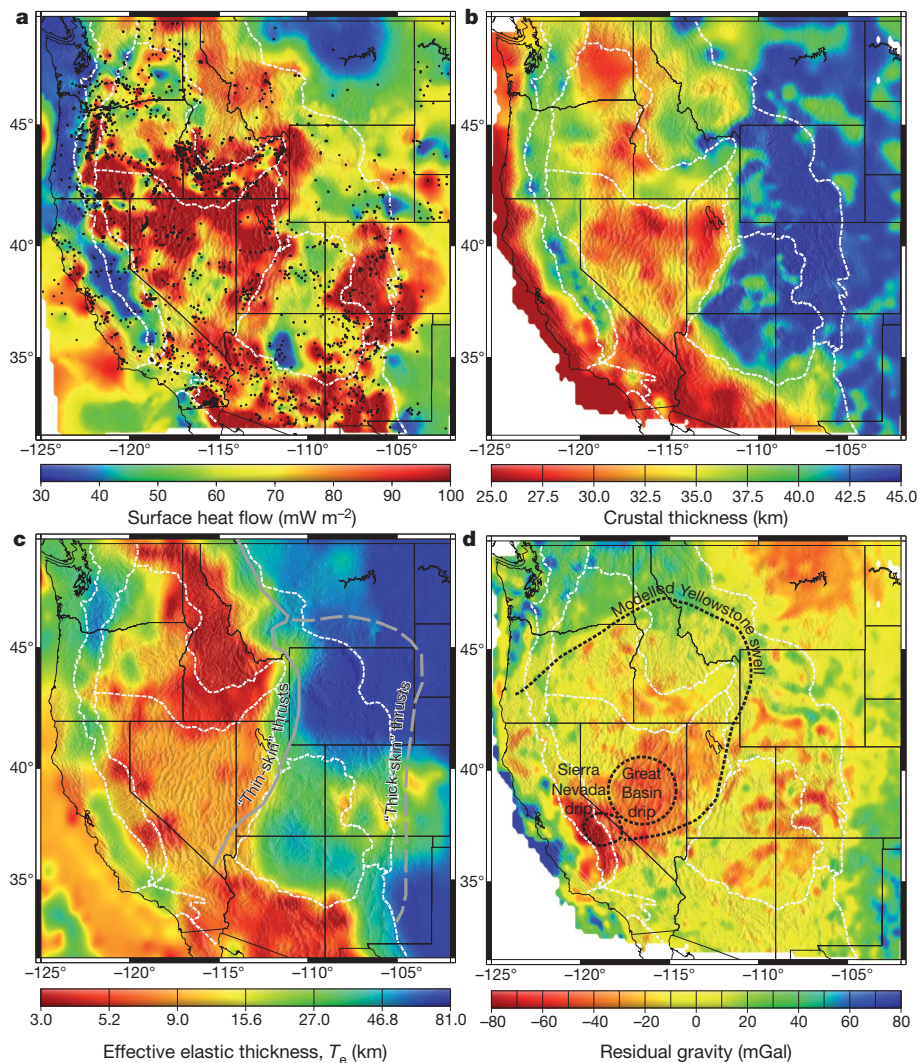


Figure 4 | Related fields. Dashed white lines are physiographic province boundaries. **a**, Surface heat flow⁷; black circles are borehole measurement sites. **b**, Crustal thickness. **c**, Effective elastic thickness from coherence analysis of gravity and topography²⁴. Grey lines approximate the eastern limits of Sevier thin-skin contraction (solid) and Laramide foreland thick-skin contraction

Also of interest is the residual of the gravity model (Fig. 4d). The residual anomalies lend confidence to the inference that unmodelled mass anomalies do not contaminate estimates of the bulk crustal composition, crustal thickness or thermal variation fields. Figure 4d shows recognizable evidence of surficial mass (for example, low density of sedimentary basins) as well as mass variation within the asthenospheric mantle. For example, residual buoyancy centred in the northern Basin and Range is similar to a ‘dynamic elevation’ anomaly previously derived from topography and crustal seismic refraction data, and interpreted as an asymmetric Yellowstone swell²⁵. However, improved resolution by this analysis suggests that return flow over seismically imaged mantle drips^{26,27} may dominate the asthenospheric mass signal. Deep-seated buoyancy anomalies such as these reflect relatively recent changes in temperature towards the base of the thermal boundary layer, and so may prove fruitful locations to look for evidence of the advective transfer processes proposed here.

In initial stages of tectonism (when geotherms are more-or-less uniform), one can reasonably expect strain focusing where crust is silica-rich and hence weak. Laramide thick-skinned thrusting in the middle Rocky Mountains of Wyoming (where modern T_e exceeds 80 km) confirms that crustal strain focusing can occur even in the presence of a thick, strong mantle lithosphere if tectonic forcing is

(dashed). **d**, Residual of the gravity model. Black dashed lines outline potassic volcanism associated with the Sierra Nevada drip²⁶, deeper imaging of the Great Basin drip²⁷, and the 500 m contour of swell elevation modelled for Yellowstone²⁵.

sufficiently robust (for example, during flat-slab subduction). One can also reasonably expect that advective heat transfer processes associated with strain will amplify lithospheric weakness. Comparisons of crustal v_p/v_s to geothermal variations suggest that this dynamical feedback mechanism plays a significant role in defining spatial distributions of tectonic strain and uplift. Examining these in the context of total lithospheric strength may improve our understanding of the roles played by the crust and mantle. EarthScope will afford an important test of the hypothesis as it moves further east and traverses the Appalachian orogeny, where Wilson first postulated cyclical repetition of tectonic events³.

METHODS SUMMARY

The analysis developed for this study introduces several innovations to the combination of complementary geophysical data. The primary data set consisted of EARS⁶ $H-K$ parameter stacks of seismic receiver-function amplitudes. Structures at crustal depths (that is, less than 50 km) are spatially aliased by near-vertical arrivals to EarthScope’s transportable array, which samples at a nominal 70 km spacing. Uncertainties inherent in a lack of redundant sampling of the crust were mitigated using spatial statistics, by deriving likelihood filters at each site from optimal interpolations (OI) of the crustal thickness and v_p/v_s estimates at neighbouring sites. A second likelihood filter was derived by modelling the Bouguer gravity field predicted by estimates at surrounding sites, allowing crustal thickness

and v_p/v_s to vary at the site of interest, and comparing to measured gravity. The combination of these filters emphasizes receiver-function stack amplitude maxima within the model-space of high likelihood rather than forcing a model parameterization preferred by the OI or gravity models. Gravity contains signals from other mass fields besides crustal thickness and crustal composition, however, and within the study region the largest is mass related to geotherm variations, or thermal mass. Thermal mass was modelled independently from geotherms estimated using surface heat flow measurements⁷ and surface radiogenic element concentrations derived from airborne γ -ray spectroscopy measurements^{28,29}. The combination of these was used to derive a model of crustal radiogenic heat production before calculation of the thermal structure. The geotherm estimates implicitly include an advective contribution to heat transfer from extensional strain²⁵; strain rates required to maintain the measured heat flow are generally two to five times larger than geodetically observed rates, suggesting a significant role for magmatic and hydrothermal processes in Cordilleran advective heat transfer. Additional methodological details are provided in Online Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 October 2010; accepted 10 February 2011.

- Williams, M. L. *et al.* Unlocking the secrets of the North American continent: an EarthScope science plan for 2010–2020. (<http://www.earthscope.org/ESSP>) (2010).
- Bürgmann, R. & Dresen, G. Rheology of the lower crust and upper mantle: evidence from rock mechanics, geodesy and field observations. *Annu. Rev. Earth Planet. Sci.* **36**, 531–567 (2008).
- Wilson, J. T. Did the Atlantic close and then re-open? *Nature* **211**, 676–681 (1966).
- Ziegler, P. A. *et al.* Dynamics of intra-plate compressional deformation: the Alpine foreland and other examples. *Tectonophysics* **252**, 7–59 (1995).
- Sykes, L. R. Intraplate seismicity, reactivation of pre-existing zones of weakness, alkaline magmatism, and other tectonism postdating continental fragmentation. *Rev. Geophys.* **16**, 621–688 (1978).
- Crotwell, H. P. & Owens, T. J. Automated receiver function processing. *Seismol. Res. Lett.* **76**, 702–709 (2005); data available at (www.iris.washington.edu/ears).
- Blackwell, D. & Richards, M. *Geothermal Map of North America, 2004* (scale 1:6,500,000, AAPG Item 423, 2004); data available at (<http://smu.edu/geothermal>).
- Christensen, N. I. Poisson's ratio and crustal seismology. *J. Geophys. Res.* **101**, 3139–3156 (1996).
- Kono, Y., Miyake, A., Ishikawa, M. & Arima, M. Temperature derivatives of elastic wave velocities in plagioclase ($An_{51\pm1}$) above and below the order-disorder transition temperature. *Am. Mineral.* **93**, 558–564 (2008).
- Roy, M., Jordan, T. H. & Pederson, J. Colorado Plateau magmatism and uplift by warming of heterogeneous lithosphere. *Nature* **459**, 978–982 (2009).
- Crossey, L. J. *et al.* Degassing of mantle-derived CO₂ and He from springs in the southern Colorado Plateau region: neotectonic connections and implications for groundwater systems. *Geol. Soc. Am. Bull.* **121**, 1034–1053 (2009).
- Zhu, L. & Kanamori, H. Moho depth variation in Southern California from teleseismic receiver functions. *J. Geophys. Res.* **105**, 2969–2980 (2000).
- Smith, R. B. & Braile, L. W. The Yellowstone hotspot. *J. Volcanol. Geotherm. Res.* **61**, 121–187 (1994).
- Trehu, A. M. *et al.* Crustal architecture of the Cascadia fore-arc. *Science* **266**, 237–243 (1994).
- Godfrey, N. J. & Klempner, S. L. Ophiolitic basement to a forearc basin and implications for continental growth: the Coast Range Great Valley ophiolite, California. *Tectonics* **17**, 558–570 (1998).
- Kistler, R. W. & Ross, D. C. A strontium isotopic study of plutons and associated rocks of the southern Sierra Nevada and vicinity, California. *Bull. US Geol. Surv.* **1920**, 1–20 (1990).
- Fleck, R. J. & Criss, R. E. Location, age, and tectonic significance of the Western Idaho Suture Zone (WISZ). *US Geol. Surv. Open File Rep.* **2004-1039**, 1–48 (2004).
- Kukkonen, I. T., Jokinen, J. & Seipold, U. Temperature and pressure dependencies of thermal transport properties of rocks: implications for uncertainties in thermal lithosphere models and new laboratory measurements of high-grade rocks in the central Fennoscandian Shield. *Surv. Geophys.* **20**, 33–59 (1999).
- Moschetti, M. P., Ritzwoller, M. H., Lin, F. & Lang, Y. Seismic evidence for widespread western-US deep-crustal deformation caused by extension. *Nature* **464**, 885–889 (2010).
- Sass, J. H. *et al.* Thermal regime of the southern Basin and Range Province 1. Heat-flow data from Arizona and the Mojave Desert of California and Nevada. *J. Geophys. Res.* **99**, 22093–22119 (1994).
- McKenzie, D. Some remarks on development of sedimentary basins. *Earth Planet. Sci. Lett.* **40**, 25–32 (1978).
- Molnar, P., England, P. & Martinod, J. Mantle dynamics, uplift of the Tibetan Plateau, and the Indian monsoon. *Rev. Geophys.* **31**, 357–396 (1993).
- Yardley, B. W. D. The role of water in the evolution of the continental crust. *J. Geol. Soc. Lond.* **166**, 585–600 (2009).
- Pérez-Gussinyé, M. *et al.* Effective elastic thickness of Africa and its relationship to other proxies for lithospheric structure and surface tectonics. *Earth Planet. Sci. Lett.* **287**, 152–167 (2009).
- Lowry, A. R., Ribe, N. M. & Smith, R. B. Dynamic elevation of the Cordillera, western United States. *J. Geophys. Res.* **105**, 23371–23390 (2000).
- Zandt, G. *et al.* Active foundering of a continental arc root beneath the southern Sierra Nevada in California. *Nature* **431**, 41–46 (2004).
- West, J. D., Fouch, M. J., Roth, J. B. & Elkins-Tanton, L. T. Vertical mantle flow associated with a lithospheric drip beneath the Great Basin. *Nature Geosci.* **2**, 439–444 (2009).
- Duval, J. S., Carson, J. M., Holman, P. B. & Darnley, A. G. Terrestrial radioactivity and gamma-ray exposure in the United States and Canada. *US Geol. Surv. Open File Rep.* **2005-1413**, (2005); available at (<http://pubs.usgs.gov/of/2005/1413/index.htm>).
- Kucks, R. P. Terrestrial radioactivity and gamma-ray exposure in the United States and Canada: Gridded geographic images. *US Geol. Surv. Open File Rep.* **2005-1413**, (2005); available at (<http://pubs.usgs.gov/of/2005/1413/datafiles.htm>).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Dueker, G. Pavlis, T. Ravat, J. Shervais and D. Schutt for discussions, and A. Braathen and R. Bürgmann for comments. We are grateful to P. Crotwell, T. Owens and members of IRIS for their efforts on the EARS database, and to P. Crotwell and B. Kucks for help with data acquisition. The work of A.R.L. on this project was supported by National Science Foundation grants EAR-0454541 (EarthScope Science: Rio Grande Rift) and EAR-0955909 (Geophysics/EarthScope CAREER: Deformation Processes), and by a Utah State University New Faculty Research Grant.

Author Contributions A.R.L. developed and implemented the joint receiver function/gravity/heatflow inversion for crustal thickness and v_p/v_s . M.P.-G. developed and implemented the inversion for effective elastic thickness. The manuscript was written by A.R.L. with contributions from M.P.-G.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.R.L. (Tony.Lowry@usu.edu).

METHODS

The joint seismic receiver function/gravity/heat flow inversion for estimating crustal thickness (H) and v_p/v_s (K) iteratively applies several data modelling and inversion procedures, summarized as a sequence of the following steps: (1) Calculate variograms of the H and K estimates. (2) Using observed Bouguer gravity data, linearly invert for optimal density parameters associated with the H and K fields plus a thermal boundary layer model derived from surface heat flow data. (3) Calculate a likelihood filter of the (H , K) parameter space at site S_i using optimal interpolation (OI) with variograms estimated in step 1. (4) Calculate a second likelihood filter at site S_i from localized gravity modelling using density parameters estimated in step 2. (5) Multiply the likelihood filters by the EARS⁶ H - K stack at site S_i , and set (H_i, K_i) to the filtered stack maximum. (6) Set $i = i + 1$ and continue from step 1.

Step 1: variogram estimation. We use raw EARS⁶ H - K parameter amplitude stacks as our receiver-function observable. Amplitude maxima from the raw stacks are assumed as the starting model of crustal thickness H and v_p/v_s ratio K . Measurement pairs are binned according to the distance between, and the variograms (that is, expected value of the squared difference as a function of distance between measurements) are estimated. Variograms of the initial (EARS⁶) and final models of H (Fig. 4b) and K (Fig. 3) are shown in Fig. 2b and c respectively. Optimal interpolation (OI, also referred to as 'kriging') uses the data variograms to interpolate data based on distance to, and estimates at, nearby measurements³⁰.

Step 2: estimation of density parameters. Modelling the gravity likelihood filter in step 4 requires an estimate of the three density parameters that relate H , K and thermal variations to gravity. Laboratory measurements and geophysical models of the Moho density contrast, $\Delta\rho_{\text{Moho}}$, the partial derivative of density with respect to v_p/v_s ratio, $\partial\rho/\partial K$, and the coefficient of thermal expansion, α_v , depend on rock composition and other factors that vary spatially, so the *in situ* values are a priori unknown. Consequently we chose to estimate a representative regional average of these parameters directly from the relationship of the fields to observed gravity within the approximate current footprint of the EarthScope Transportable Array.

We first calculate Bouguer gravity fields per unit density, B_H^1 , B_K^1 and B_T^1 , by neglecting the density parameters $\Delta\rho_{\text{Moho}}$, $\partial\rho/\partial K$ and α_v . We reference all estimates of crustal thickness H to a datum (that is, subtract the site elevations) and then optimally interpolate H and K to a 16-km mesh grid covering the region from 32–49° N latitude, 102–125° E longitude. Means \bar{H} and \bar{K} are then removed, and the fields are mirrored and Fourier-transformed to frequency domain amplitudes $\tilde{H} = F\{H - \bar{H}\}$ and $\tilde{K} = F\{K - \bar{K}\}$, where $F\{\cdot\}$ denotes the Fourier transform operator. We calculate Fourier amplitudes of the Bouguer per unit Moho density contrast, \tilde{B}_H^1 , associated with the H field via:

$$\tilde{B}_H^1 = 2\pi G \tilde{H} \exp(-k\bar{H}) \quad (1)$$

where G is the universal gravitational constant and k is modulus of the two-dimensional wavenumber associated with each Fourier amplitude. The Bouguer per unit density amplitudes, \tilde{B}_K^1 , associated with K are calculated via:

$$\tilde{B}_K^1 = 2\pi G \left[\frac{1 - \exp(-k\bar{H})}{k} \tilde{K} - \tilde{M} \exp(-k\bar{H}) \right] \quad (2)$$

in which $\tilde{M} = F\{(\bar{H} - \bar{K})(K - \bar{K})\}$ is a correction factor for variable crustal thickness.

Because thermal variations contribute significantly to the total gravity used here to invert for crustal thickness and composition, a correction for estimated thermal mass is included in the model. The model of thermal gravity uses a geotherm inverted from data that include surface heat flow measurements⁷, q_s , airborne γ -ray spectroscopy measurements of surface radiogenic element concentrations^{28,29}, and mean annual surface temperature data³¹. Supplementary Fig. 1a depicts the OI of heat flow measurements⁷. These were extrapolated to depth using one-dimensional heat-transfer relations that assume steady-state advection by extensional strain²⁵, and a thermal conductivity, κ , with temperature-dependent parameterization³² that averages crustal and mantle rock types¹⁶. We assume mantle potential temperature $T_r = 1,320^\circ\text{C}$, and an adiabat is derived for pressure/temperature-dependent values of mantle coefficient of thermal expansion and specific heat capacity³³. The airborne γ -ray spectroscopy measurements^{28,29} of U, Th and K radiogenic element concentrations (Supplementary Fig. 1c) are converted to surface heat production, A_0 , (Supplementary Fig. 1d) via³⁴:

$$A_0 = \rho(3.58 \times 10^{-5} C_K + 2.69 \times 10^{-5} C_{\text{Th}} + 9.71 \times 10^{-5} C_U) \quad (3)$$

in which C_X are concentrations in parts per billion. Weighted least squares regression of the relationship between q_s and A_0 within 500 km windows yields estimates of a length-scale for an exponential depth decay, l_{rad} , that are physically reasonable (Supplementary Fig. 1e); a posteriori (that is, misfit-scaled) parameter uncertainties of radiogenic length estimates are typically 10–30%. In the absence of better constraint, this approach adequately represents the spatial distribution of heat producing elements.

After subtracting the portion arising from shallow heat production, heat flow is filtered³⁵ to remove effects of sources and sinks shallower than 30 km (Supplementary Fig. 1b). This removes most (though not all) advective effects of near-surface hydrology and heat flow measurement noise associated with three-dimensional conductivity variations and topographic effects. Temperature is estimated as²⁵:

$$T(z) = T_s + \frac{A_0 l_{\text{rad}}^2}{\kappa} \left[1 - \exp\left(-\frac{z}{l_{\text{rad}}}\right) \right] + \left(T_r - \frac{A_0 l_{\text{rad}}^2}{\kappa} - T_s \right) \text{erf}\left(\frac{z}{l_{\text{con}}}\right) \quad (4)$$

where

$$l_{\text{con}} = \frac{2[\kappa(T_r - T_s) - A_0 l_{\text{rad}}^2]}{\sqrt{\pi}(q_s - A_0 l_{\text{rad}})} \quad (5)$$

and the minimum and maximum estimates of geotherms from the map region are shown in Supplementary Fig. 1c inset. (Note the actual relations used are complicated slightly, relative to equations (4) and (5), by the temperature-dependence of κ and a solution for two relations, one each for the crust and mantle, with stipulation that temperature and heat flux for both are identical at the Moho).

Supplementary Fig. 1f shows the surface temperature, T_s , adapted from climate models of mean annual temperature³¹. An advective contribution from extensional strain (needed to accommodate advective heat transfer) is implicit rather than explicit in these relations. Map-view thermal fields, T_z , are averaged over 1-km depth ranges for $z = 1, \dots, 200$ km, means are removed, and the data are mirrored and transformed as $\tilde{T}_z = F\{T_z - \bar{T}_z\}$. Then amplitudes of the thermal gravity field per unit density variation, \tilde{B}_T^1 , are calculated via:

$$\tilde{B}_T^1 = 2,000\pi G \sum_{z=1}^{200} \bar{\rho}_z \tilde{T}_z \exp(-kz) \quad (6)$$

Here $\bar{\rho}_z$ is an expected value for density at a given depth, based on the mean temperature and a continental mean compositional profile with depth³⁶. Note the multiplication by factor-of-1,000 arises from the 1-km discretisation of the geotherm.

We now have amplitudes for three different gravity fields, B_H^1 , B_K^1 and B_T^1 , that are independent of density parameters. To estimate those density parameters, we inverse transform the amplitudes to the spatial domain, strip mirrored portions and then invert the highly over-determined system of linear equations:

$$\tilde{B}_H^1 \Delta\rho_{\text{Moho}} + \tilde{B}_K^1 \frac{\partial\rho}{\partial K} + \tilde{B}_T^1 \alpha_v + \tilde{s} = \tilde{B}_{\text{obs}} \quad (7)$$

for the unknown parameters $\Delta\rho_{\text{Moho}}$, $\partial\rho/\partial K$, α_v , and a static offset s . Here, the arrow denotes a $\sim 15,000 \times 1$ vector.

Supplementary Fig. 2 depicts each of the individual components of the final regional model of the gravity field (Supplementary Fig. 2a–c), the observed-minus-modelled residual (Supplementary Fig. 2d; see also Fig. 4d), the observed gravity field (Supplementary Fig. 2e), and the total modelled gravity field summing contributions from the Moho, crustal density variations and thermal density variations (Supplementary Fig. 2f). It is worth noting that the mass fields derived from this gravity modelling can be used for geodynamical investigations of stress, strain and isostatic loading, and in fact estimation of these mass fields was the original objective of this investigation. Supplementary Fig. 3 depicts evolution of the r.m.s. gravity residual for the summed model, and for separate components of the model, over $\sim 25,000$ site updates (roughly 16 iterations over all seismic sites within the map space). Residual of the summed model drops from 55 mGal for the starting model, using raw EARS estimates of H and K , to 26 mGal for the final model. By comparison, the root-variance of observed gravity is 68 mGal. (Note that the deviation beginning around 15,000 site updates resulted after introduction of new H - K stacks from the IRIS database⁶. The iterative inversion results shown here required about six months of computation on a quad-core workstation, during which time the EarthScope Transportable Array continued to collect new earthquake data and add new sites). Taken separately, the raw EARS data contribute virtually no reduction of variance to the starting model. Hence, the thermal model dominates total reduction of variance when the raw EARS estimates are used, that is, at iteration 0 in Supplementary Fig. 3. The temperature model does not change during the inversion run (although the thermal density parameter, α_v , does), and contributions from the Moho and crustal density models overtake the thermal mass contribution by the end of the second iteration over all sites (with crustal density eventually dominating the overall reduction of gravity variance).

The final density parameter estimates are $0.81 \times 10^{-5} ^\circ\text{C}^{-1}$ for the coefficient of thermal expansion α_v , 115 kg m^{-3} for $\Delta\rho_{\text{Moho}}$, and 460 kg m^{-3} for $\partial\rho/\partial K$. These values differ slightly from expectations based on laboratory experiments and other geophysical studies. Laboratory data on pressure- and temperature-dependence of the thermal expansion coefficient for rocks within the thermal boundary layer predict values in the range $(0.5\text{--}3.5) \times 10^{-5} ^\circ\text{C}^{-1}$. The low estimate derived here is consistent with predictions for granite³⁷, but the upper mantle (where α_v should be

$(2.5\text{--}3.5) \times 10^{-5}$; ref. 38) dominates the thermal mass integral (Supplementary Fig. 1c inset). Recent studies suggest the effective thermal expansivity of a viscoelastic fluid mantle will be 15–30% less than laboratory mineral physics measurements of the volumetric coefficient of thermal expansion^{39,40}, but the ~70% lower value estimated here is more likely to reflect a poor geotherm model. Surface heat flow measurements are sparse and notoriously noisy, and non-steady-state contributions to lithospheric heat transfer (for example, under the Sierra Nevada⁴¹ and the Colorado plateau²³) cannot be modelled accurately from surface data alone, so the thermal model used here is likely to represent mantle temperature variations poorly. One implication of this is that the true correlation coefficient of geothermal variations and crustal composition may be even higher than we have estimated.

The density contrast for continental Moho inferred from normal mode seismic analysis is 480 kg m^{-3} (ref. 42), but laboratory measurements indicate it can range from about -200 to 600 kg m^{-3} and globally averages $\sim 230 \text{ kg m}^{-3}$ (ref. 36). Our final crustal thickness estimates (Fig. 4b) match fairly well the independent estimates from more sparsely sampled crustal seismic refraction studies (for example, compiled in ref. 25) and from inversion of Rayleigh phase velocity data (D. Schutt, personal communication). The regression slope of the latter independent estimate with ours is 1.001 with a correlation coefficient of 0.72, so the density contrast estimate is unlikely to be significantly in error. Much of the thicker crust in western portions of the region of our inversion is underlain by dense cumulates and/or eclogites, for example, under the Sierra Nevada²⁶, Cascade Range and the Snake River plain⁴³; other regions of thick crust (such as the Wyoming province⁴⁴) exhibit anomalously high lower crustal velocity. These would imply a lower-than-average Moho density contrast. Laboratory measurements suggest the partial derivative of density with respect to v_p/v_s deriving from compositional variations should be around $1,500 \text{ kg m}^{-3}$ (Fig. 1a), although this will vary significantly depending on actual abundances of various compositions, and the direct effect of temperature on v_p/v_s effectively would lower the observational slope.

Step 3: OI likelihood filter. OI yields both an expected value and an estimate of uncertainty, and we use it in this application both for interpolation of various measurement fields to arbitrary locations and for estimation of likelihood functions applied to H - K stacks. The OI likelihood filter at a seismic site S_i chosen for update is estimated by interpolating H and K estimates from the 150 nearest sites to the location of site S_i . The expected values (H_i^{OI} , K_i^{OI}) and 1σ estimates of uncertainty (σ_H , σ_K) are used to determine OI confidence interval C_{OI} at all stack amplitude realizations (H_a , K_a) present in the EARS stack via:

$$C_{\text{OI}}(H_a, K_a) = \left(\frac{H_a - H_i^{\text{OI}}}{\sigma_H} \right)^2 + \left(\frac{K_a - K_i^{\text{OI}}}{\sigma_K} \right)^2 \quad (8)$$

Then, assuming H and K are Gaussian processes, C_{OI} will be χ^2 -distributed and the likelihood corresponding to that confidence interval is given by⁴⁵:

$$L = \frac{1}{\sqrt{1 + \frac{M}{N-M} F_{\alpha}^{-1}(M, N-M)}} \quad (9)$$

Here the degrees of freedom are those of the H - K stack: M , the number of model parameters, is 2 (being H and K); the number of observations $N = 4n$ where n is the number of earthquakes used in the stack and 4 represents the number of seismic phases sampled (Ps, PpSs, PsPs and PpPs), F^{-1} is the inverse of the F -cumulative distribution function, and α represents the probability:

$$\alpha = \exp\left(-\frac{C_{\text{OI}}}{2}\right) \quad (10)$$

An example OI likelihood filter $L(H, K)$ at Nevada site TA.O09A is shown in Supplementary Fig. 4a.

Step 4: gravity likelihood filter. A second likelihood filter is derived from gravity modelling within a relatively small window near the seismic site. For each possible permutation of (H_a , K_a) represented in the EARS amplitude stack, we temporarily set (H_i , K_i) at the update site S_i equal to (H_a , K_a) and hold fixed the crustal thickness and v_p/v_s ratio at surrounding sites. We reference crustal thickness H to a datum (that is, subtract the site elevation) and then optimally interpolate H and K to a 9-km mesh, 288-km aperture grid centred at site S_i . The gravity modelling is analogous to that used to estimate density parameters in step 2, except that now we model Bouguer gravity (that is, using the estimated density parameters) rather than gravity per unit density. Means \bar{H} and \bar{K} are removed, the fields are mirrored and Fourier-transformed to frequency domain amplitudes $\tilde{H} = F\{H - \bar{H}\}$ and $\tilde{K} = F\{K - \bar{K}\}$, and we calculate amplitudes of the Bouguer gravity fields \tilde{B}_H associated with the H field via:

$$\tilde{B}_H = 2\pi G \Delta \rho_{\text{Moho}} \tilde{H} \exp(-k\bar{H}) \quad (11)$$

\tilde{B}_K associated with K via:

$$\tilde{B}_K = 2\pi G \frac{\partial \rho}{\partial K} \left[\frac{1 - \exp(-k\bar{H})}{k} \tilde{K} - \tilde{M} \exp(-k\bar{H}) \right] \quad (12)$$

and \tilde{B}_T for the thermal model as:

$$\tilde{B}_T = 2,000\pi G \alpha_v \sum_{z=1}^{200} \bar{\rho}_z \tilde{T}_z \exp(-kz) \quad (13)$$

We inverse transform gravity fields to the spatial domain, strip mirrored portions and calculate a modelled gravity field B_{Mod} by summing B_{Th} , B_K and B_T . Example observed and modelled gravity for the TA.O09A site is shown in Supplementary Fig. 5a and b, respectively. We then subtract the mean difference between modelled and observed gravity, and calculate the L_2 -norm of the residual difference (Supplementary Fig. 5c). The r.m.s. misfit, R , as a function of assumed (H_a , K_a) at the site (Supplementary Fig. 5d) can be used to derive confidence intervals on the gravity model via the likelihood ratio method⁴⁵, in which the confidence region with probability α of containing the correct solution corresponds to the model space for which

$$R^2 \leq R_{\min}^2 \left[1 + \frac{M}{N_g - M} F_{\alpha}^{-1}(M, N_g - M) \right] \quad (14)$$

where N_g is the number of gravity observations and R_{\min} is the global minimum gravity residual. We map R^2 to given probability α and then use α as a function of H - K to calculate a likelihood filter for the stack amplitude, using equation (9). This effectively transforms the misfit function from the gravity modelling to the degrees-of-freedom inherent to the receiver function stack. An example likelihood filter derived from gravity modelling at site TA.O09A is shown in Supplementary Fig. 4b. **Step 5: updating the H - K estimate at a site.** Both the OI and gravity likelihood filters (Supplementary Fig. 4c) are multiplied by the raw EARS receiver function amplitude stack (Fig. 2d). Then (H_a , K_a) corresponding to the amplitude maximum of the filtered stack (Fig. 2e and Supplementary Fig. 4d) are adopted as the updated estimate of (H_i , K_i) at site S_i . These updates are performed iteratively over all sites within the map region.

Likelihood filters for this analysis do not force estimates to match the variogram and gravity constraints, but instead guide the model away from receiver function H - K stack maxima that result from interference with structural complexities. Supplementary Fig. 6 shows observed receiver functions (time-scaled to remove dependence of Ps arrival time on ray parameter) for the 53 events recorded at site TA.O09A. These, and their stack, are compared to a synthetic receiver function generated for the simple Earth model ($H = 29.75$; $K = 1.76$) derived by our method, demonstrating that results reported here are consistent with the raw receiver function data.

30. Davis, J. C. *Statistics and Data Analysis in Geology* 2nd edn (Wiley, 1986).
31. Grieser, J., Gommers, R., Cofield, S. & Bernardi, M. Data sources for FAO worldmaps of Koeppen climatologies and climatic net primary production. (http://www.fao.org/nr/climpag/globgrids/KC_commondata_en.asp) (2006).
32. Seipold, U. Temperature dependence of thermal transport properties of crystalline rocks — a general law. *Tectonophysics* **291**, 161–171 (1998).
33. Bouhifd, M. A., Andrault, D., Fiquet, G. & Richet, P. Thermal expansion of forsterite up to the melting point. *Geophys. Res. Lett.* **23**, 1143–1146 (1996).
34. Turcotte, D. L. & Schubert, G. *Geodynamics* 2nd edn (Cambridge Univ. Press, 2002).
35. Mareschal, J. C., Cunningham, J. P. & Lowell, R. P. Downward continuation of heat flow data: method and examples from the western United States. *Geophysics* **50**, 846–851 (1985).
36. Christensen, N. I. & Mooney, W. D. Seismic velocity structure and composition of the continental crust: a global view. *J. Geophys. Res.* **100**, 9761–9788 (1995).
37. Heard, H. C. & Page, L. Elastic moduli, thermal expansion and inferred permeability of two granites to 350 °C and 55 megapascals. *J. Geophys. Res.* **87**, 9340–9348 (1982).
38. Afonso, J. C., Ranalli, G. & Fernández, M. Thermal expansivity and elastic properties of the lithospheric mantle: results from mineral physics of composites. *Phys. Earth Planet. Inter.* **149**, 279–306 (2005).
39. Pollack, H. N. On the use of the volumetric thermal expansion coefficient in models of ocean floor topography. *Tectonophysics* **64**, T45–T47 (1980).
40. Korenaga, J. Effective thermal expansivity of Maxwellian ocean lithosphere. *Earth Planet. Sci. Lett.* **257**, 343–349 (2007).
41. Erkan, K. & Blackwell, D. Transient thermal regimes in the Sierra Nevada and Baja California extinct outer arcs following the cessation of Farallon subduction. *J. Geophys. Res.* **114**, doi:10.1029/2007JB005498 (2009).
42. Dziewonski, A. M. & Anderson, D. L. Preliminary Reference Earth Model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
43. DeNosaquo, K. R., Smith, R. B. & Lowry, A. R. Density and lithospheric strength models of the Yellowstone-Snake River Plain volcanic system from gravity and heat flow data. *J. Volcanol. Geotherm. Res.* **188**, 108–127 (2009).
44. Yuan, H. Y., Dueker, K. & Stachnik, J. Crustal structure and thickness along the Yellowstone hot spot track: evidence for lower crustal outflow from beneath the eastern Snake River Plain. *Geochem. Geophys. Geosyst.* **11**, doi:10.1029/2009GC002787 (2010).
45. Beck, J. V. & Arnold, K. J. *Parameter Estimation in Engineering and Science* (Wiley, 1977).

Amygdala circuitry mediating reversible and bidirectional control of anxiety

Kay M. Tye^{1*}, Rohit Prakash^{1,2*}, Sung-Yon Kim^{1,2*}, Lief E. Fenno^{1,2*}, Logan Grosenick^{1,2}, Hosniya Zarabi¹, Kimberly R. Thompson¹, Viviana Gradinaru^{1,2}, Charu Ramakrishnan¹ & Karl Deisseroth^{1,2,3,4,5}

Anxiety—a sustained state of heightened apprehension in the absence of immediate threat—becomes severely debilitating in disease states¹. Anxiety disorders represent the most common of psychiatric diseases (28% lifetime prevalence)² and contribute to the aetiology of major depression and substance abuse^{3,4}. Although it has been proposed that the amygdala, a brain region important for emotional processing^{5–8}, has a role in anxiety^{9–13}, the neural mechanisms that control anxiety remain unclear. Here we explore the neural circuits underlying anxiety-related behaviours by using optogenetics with two-photon microscopy, anxiety assays in freely moving mice, and electrophysiology. With the capability of optogenetics^{14–16} to control not only cell types but also specific connections between cells, we observed that temporally precise optogenetic stimulation of basolateral amygdala (BLA) terminals in the central nucleus of the amygdala (CeA)—achieved by viral transduction of the BLA with a codon-optimized channelrhodopsin followed by restricted illumination in the downstream CeA—exerted an acute, reversible anxiolytic effect. Conversely, selective optogenetic inhibition of the same projection with a third-generation halorhodopsin¹⁵ (eNpHR3.0) increased anxiety-related behaviours. Importantly, these effects were not observed with direct optogenetic control of BLA somata, possibly owing to recruitment of antagonistic downstream structures. Together, these results implicate specific BLA–CeA projections as critical circuit elements for acute anxiety control in the mammalian brain, and demonstrate the importance of optogenetically targeting defined projections, beyond simply targeting cell types, in the study of circuit function relevant to neuropsychiatric disease.

Despite the high prevalence^{1,2} of anxiety disorders, the underlying neural circuitry is incompletely understood. Available treatments are inconsistently effective or, in the case of benzodiazepines, addictive and linked to significant side effects including cognitive impairment and respiratory suppression¹⁷, pointing to the need for a deeper understanding of anxiety control mechanisms in the mammalian brain.

Although amygdala microcircuitry for conditioned fear has been optogenetically dissected^{18,19}, the causal underpinnings of unconditioned anxiety¹¹ have not yet been investigated with cellular precision. Pointing to the need for precise optogenetic exploration, the amygdala is composed of functionally and morphologically heterogeneous subnuclei with complex interconnectivity. The BLA is primarily glutamatergic (~90%)^{20,21} whereas the CeA, which encompasses the centrolateral (CeL) and centromedial (CeM) nuclei, consists of ~95% GABAergic medium spiny neurons²². The primary output region of the amygdala is the CeM^{23,24}, which (when chemically or electrically excited) mediates autonomic and behavioural responses associated with fear and anxiety via projections to the brainstem²⁵. Because patients with generalized anxiety disorder may have abnormal activity arising from the BLA and CeM¹¹, and as BLA neurons excite GABAergic CeL neurons²⁶ that

provide feed-forward inhibition onto CeM ‘output’ neurons^{6,18}, we considered that the BLA–CeL–CeM circuitry could be causally involved in anxiety. However, BLA pyramidal neurons as a whole could have varied and antagonistic roles in diverse projections throughout the brain, with targets including the bed nucleus of the stria terminalis (BNST), nucleus accumbens, hippocampus and cortex²⁶.

We therefore developed a method to selectively control BLA terminals in the CeA (Supplementary Methods). BLA glutamatergic projection neurons were transduced with an adeno-associated virus serotype 5 (AAV5) carrying codon-optimized channelrhodopsin (ChR2)–eYFP under control of the CaMKII α promoter followed by unilateral implantation of a bevelled guide cannula to allow preferential illumination of the non-transduced CeL (Supplementary Figs 1, 2). *In vivo* electrophysiological recordings were used to determine illumination parameters for selective control of those BLA terminals in the CeA without nonspecific control of all BLA somata (Supplementary Fig. 3).

To investigate the functional role of the BLA–CeA pathway in anxiety, we probed freely moving mice under projection-specific optogenetic control in two well-validated²⁷ anxiety assays: the elevated-plus maze (EPM) and the open-field test (OFT) (Fig. 1a–e). Mice display anxiety-related behaviours in open spaces; therefore, increased time spent in the open arms of the EPM or in the centre of the OFT is interpreted as reduced anxiety²⁷. To test whether anxiety-related behaviours could be related to activation of the BLA–CeA projection, and not all BLA somata as a whole, we compared mice receiving projection-specific photostimulation (ChR2:BLA–CeA; Fig. 1a) to a group with identical illumination parameters transduced with a control virus (eYFP:BLA–CeA), and to another control group expressing ChR2 in the BLA receiving direct illumination of the BLA (ChR2:BLA(somata)). Photostimulation of BLA terminals in the CeA (ChR2:BLA–CeA) increased open-arm time ($F_{1,42} = 69.09$, $P < 0.00001$; Fig. 1b, c) and the probability of open-arm entry from the maze centre ($F_{1,42} = 24.69$, $P < 0.00001$; Fig. 1c, inset, and Supplementary Movie) on the EPM, as well as increasing centre time in the OFT ($F_{1,105} = 24.46$, $P < 0.00001$; Fig. 1d, e), reflecting anxiety reduction without impaired locomotion (Supplementary Fig. 4). In contrast, the ChR2:BLA(somata) group showed reduced open-arm time ($F_{1,42} = 6.20$, $P = 0.02$; Fig. 1b, c) and probability of open-arm entry ($F_{1,42} = 5.15$, $P = 0.03$) during photostimulation relative to eYFP:BLA–CeA controls, reflecting a distinct anxiogenic effect. Thus, selective illumination of BLA projections to the CeA, but not of BLA somata nonspecifically, produced an acute, rapidly reversible anxiolytic effect.

Next we investigated the physiological basis of this light-induced anxiolysis. We considered that preferential photostimulation of BLA terminals in the CeL could activate CeL neurons and exert feed-forward inhibition onto brainstem-projecting CeM output neurons¹⁸ to implement anxiolysis. To test this, we undertook *in vivo* experiments, with light-delivery protocols matched to those delivered in the behavioural experiments, using activity-dependent immediate early

¹Department of Bioengineering, Stanford University, Stanford, California 94305, USA. ²Neurosciences Program, Stanford University, Stanford, California 94305, USA. ³Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California 94305, USA. ⁴Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA. ⁵CNC Program, Stanford University, Stanford, California 94305, USA.

*These authors contributed equally to this work.

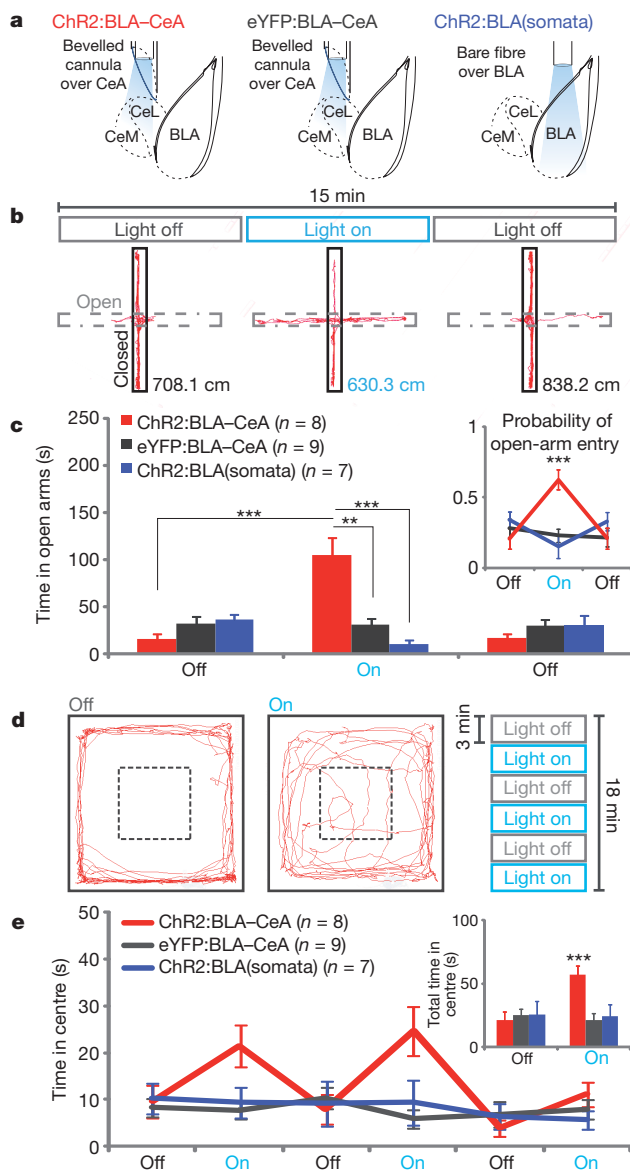


Figure 1 | Projection-specific excitation of BLA terminals in the CeA induces acute reversible anxiolysis. **a**, Mice were housed in a high-stress environment before behavioural manipulations, and received 5-ms light pulses at 20 Hz for all light-on conditions. **b**, **c**, ChR2:BLA-CeA mice ($n = 8$) received selective illumination of BLA terminals in the CeA during the light-on epoch on the EPM; see ChR2:BLA-CeA representative path (**b**), which induced an increase in open-arm time on photostimulation relative to eYFP:BLA-CeA ($n = 9$) and ChR2:BLA(somata) ($n = 7$) controls (**c**), and an increase in the probability of open-arm entry (see inset). **d**, **e**, ChR2:BLA-CeA mice also showed increased centre time on the OFT, as seen in a representative path (**d**), during light-on epochs relative to light-off epochs and eYFP:BLA-CeA and ChR2:BLA(somata) controls (**e**).

gene (*c-fos*) expression to track neuronal activation. We quantified the proportion of neurons in the BLA, CeL and CeM (Supplementary Fig. 5) within each group that expressed eYFP or showed *c-fos* immunoreactivity. Opsin expression was specific to BLA glutamatergic neurons, and was not observed in intercalated cells (Supplementary Fig. 6). No group differences were detected in the proportion of eYFP-positive cells within each region (Supplementary Fig. 5). We found a significantly higher proportion of *c-fos*-positive BLA cells in the ChR2:BLA(somata) group ($F_{2,9} = 10.12$, $P < 0.01$), relative to the ChR2:BLA-CeA or eYFP:BLA-CeA groups ($P < 0.01$ and $P < 0.05$, respectively), but no detectable difference between the ChR2:BLA-CeA and eYFP:BLA-CeA groups, indicating that the bevelled cannula shielding effectively

prevented BLA somata photostimulation. A higher proportion of CeL neurons expressed *c-fos* in the ChR2:BLA-CeA group relative to the eYFP:BLA-CeA group ($F_{2,9} = 4.54$, $P = 0.04$), but not the ChR2:BLA(somata) group (Supplementary Fig. 5). Thus, the *in vivo* illumination of BLA-CeA projections that triggered acute anxiolysis was found to excite CeL neurons without activating BLA somata.

To test the hypothesis that preferential illumination of BLA-CeL terminals induced feed-forward inhibition of CeM output neurons, we combined whole-cell patch-clamp electrophysiology with two-photon imaging to visualize the microcircuit while simultaneously probing the functional relationships among these cells during projection-specific optogenetic control (Fig. 2a–f). BLA neurons showed high-fidelity spiking to direct illumination (Fig. 2b). Illumination of BLA terminals in CeL elicited excitatory responses with stable spiking fidelity, and responding cells included both weakly and strongly excited CeL cells ($n = 16$; Fig. 2c).

To test whether illumination of BLA-CeL synapses could block CeM spiking via feed-forward inhibition from CeL neurons, we recorded from CeM neurons while selectively illuminating BLA-CeL synapses (Fig. 2d). Indeed, we observed potent spiking inhibition in CeM cells on illumination of BLA-CeL terminals (Fig. 2d; $F_{2,11} = 15.35$, $P = 0.004$). Figure 2e shows CeM responses recorded during illumination of ChR2-expressing BLA neurons in the CeM; importantly, the very same CeM neurons ($n = 7$) showed net excitation with broad illumination of BLA inputs to the CeM (Fig. 2e), but showed net inhibition on selective illumination of BLA inputs to the CeL (Fig. 2f). These data from a structurally and functionally identified microcircuit²⁵ illustrate that the balance of direct and indirect inputs from the BLA to the CeM can modulate CeM activity. We then directly tested whether overlapping or distinct populations of BLA neurons projected to the CeL and CeM in the mouse (Supplementary Fig. 7) by two-photon imaging in 350- μ m-thick coronal slices. Of the BLA neurons sampled ($n = 18$; Supplementary Fig. 7), 44% projected to the CeL alone and 17% projected to the CeM alone, with only one cell observed to project to both the CeL and the CeM.

Because *in vivo* *c-fos* assays had indicated that illumination of BLA terminals in the CeL sufficed to drive excitation of postsynaptic CeL neurons but not to recruit efficiently BLA neurons as a whole, we next investigated the properties of optogenetically driven terminal stimulation in this microcircuit using whole-cell recordings. We first recorded from BLA pyramidal neurons expressing ChR2 and moved a restricted light spot (~ 125 μ m in diameter) in 100- μ m steps from the cell soma, both in a direction over the visually identified axon and in a direction where there was no axon collateral, illustrating the spatial properties of light scattering in this circuit (Fig. 3a, b); spiking fidelity in the BLA neuron and evoked inward currents are summarized (Fig. 3a). Next, we found that typical photostimulation parameters drove reliable transmission when delivered to BLA-CeA synapses (assayed with recordings from postsynaptic CeL neurons; Fig. 3b and Supplementary Fig. 8); in contrast, when recording from the BLA somata instead, ~ 300 μ m from the light spot, we did not observe reliable antidromically driven action potential firing (only $\sim 5\%$ reliability) despite use of the very same BLA-CeA synapse illumination conditions that elicited 100% reliable transmitter release from the illuminated terminals and the same cells known to spike robustly in response to somatic illumination (Fig. 3b). These results were consistent with the *c-fos* immunoreactivity and behavioural data (Fig. 1) and held even on bath application of GABA and glutamate receptor antagonists ($n = 7$) to eliminate local circuitry effects. The marked difference between effective synaptic transmission and antidromic spiking fidelity ($P = 0.0039$; Fig. 3b, inset) reveals that optogenetically driven vesicle release can occur in the absence of reliable antidromic drive, a potentially useful property that may relate to projection parameters such as axon calibre and myelination status (optogenetic stimuli will recruit thinner axons more efficiently than electrical stimuli), as well as experimental light intensity and spatial restriction properties.

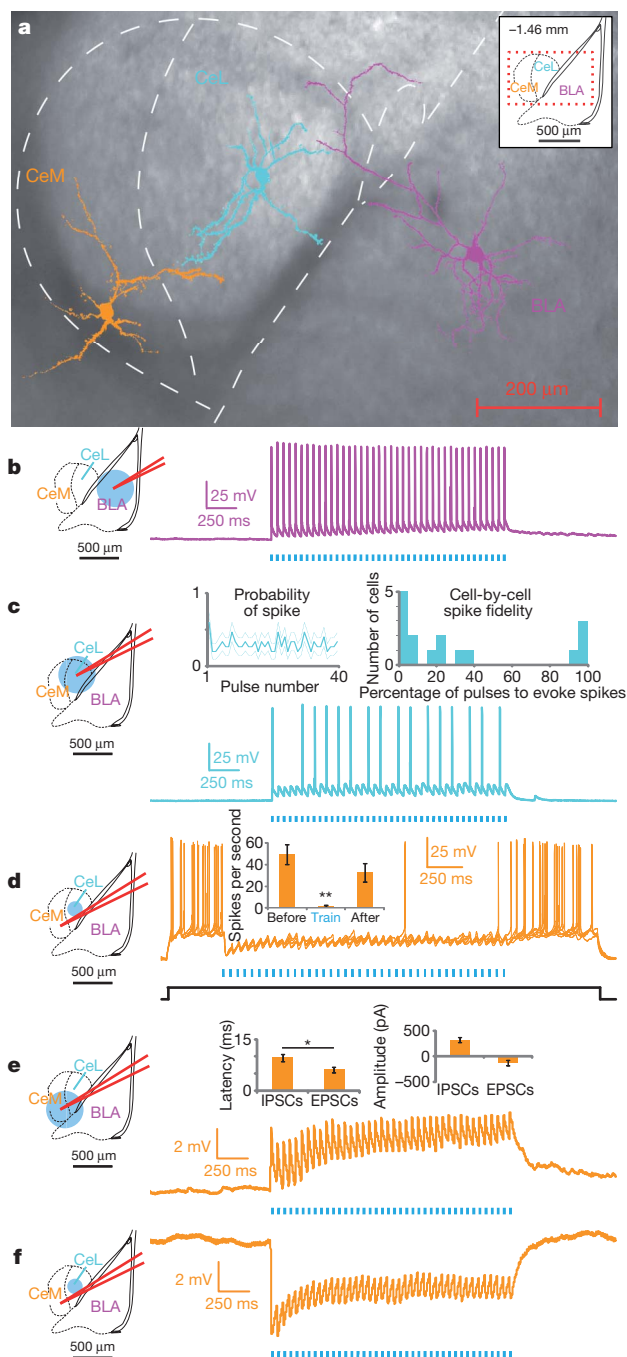


Figure 2 | Projection-specific excitation of BLA terminals in the CeA activates CeL neurons and elicits feed-forward inhibition of CeM neurons.

a, Two-photon images of representative BLA, CeL and CeM cells imaged from the same slice, overlaid on a brightfield image. **b–f**, Schematics of the recording and illumination sites for the associated representative current-clamp traces (membrane potential $V_m \sim -70$ mV). **b**, Representative BLA pyramidal neuron trace expressing ChR2, all of which spiked for every pulse ($n = 4$). **c**, Representative trace from a CeL neuron in the terminal field of BLA projection neurons, showing both sub- and suprathreshold excitatory responses on photostimulation ($n = 16$). Inset left, population summary of mean probability of spiking for each pulse in a 40-pulse train at 20 Hz, dotted lines indicate s.e.m. Inset right, frequency histogram showing individual cell spiking fidelity; y-axis is the number of cells per each 5% bin. **d**, Six sweeps from a CeM neuron spiking in response to a current step (~ 60 pA; indicated in black) and inhibition of spiking on 20 Hz illumination of BLA terminals in the CeL. Inset, spike frequency was significantly reduced during light stimulation of CeL neurons ($n = 4$; spikes per second before (49 ± 9.0), during (1.5 ± 0.87) and after (33 ± 8.4) illumination; mean \pm s.e.m.). **e, f**, On broad illumination of the CeM, voltage-clamp summaries show that the latency of excitatory postsynaptic currents (EPSCs) is significantly shorter than the latency of inhibitory postsynaptic currents (IPSCs), whereas there was a non-significant difference in the amplitude of EPSCs and IPSCs ($n = 11$; $*P = 0.04$, see insets). The same CeM neurons ($n = 7$) showed either net excitation when receiving illumination of the CeM (**e**) or net inhibition on selective illumination of the CeL (**f**).

drug treatment did not impair locomotor activity (Supplementary Fig. 9), and in acute slices time-locked light-evoked excitatory responses were abolished on bath application of 2,3-dihydroxy-6-nitro-7-sulfamoyl-benzo[f]quinoxaline-2,3-dione (NBQX) and (2R)-amino-5-phosphonovaleric acid; (2R)-amino-5-phosphonopentanoate (AP5) (Supplementary Fig. 10). These data demonstrate that the light-induced anxiolytic effects were caused by the activation of BLA–CeA synapses, and not attributable to BLA projections to distal targets passing through the CeA.

Finally, to test whether basal anxiety-reducing processes could be blocked by selectively inhibiting the BLA–CeA pathway, we bilaterally transduced either eNpHR3.0—which hyperpolarizes neuronal membranes on illumination with amber light¹⁵—or eYFP alone, under the CaMKII α promoter in the BLA, and implanted bilateral bevelled guide cannulae to allow selective illumination of BLA terminals in the CeA (Supplementary Fig. 11). eNpHR3.0 expression was restricted to glutamatergic neurons in the BLA, and the eNpHR3.0:BLA–CeA group showed raised levels of *c-fos* expression relative to the eYFP:BLA–CeA and eNpHR3.0:BLA(somata) groups in the CeM ($P < 0.05$; Supplementary Fig. 12), consistent with the hypothesis that inhibition of BLA–CeL synapses suppresses feed-forward inhibition from CeL neurons to CeM neurons, thereby increasing CeM excitability and the downstream processes leading to increased anxiety phenotypes. Selective illumination of eNpHR3.0-expressing axon terminals reduced the probability of both spontaneously occurring (frequency: $F_{1,8} = 32.99$, $P = 0.00024$; amplitude: $F_{1,8} = 21.96$, $P = 0.001$; Supplementary Fig. 13) and electrically evoked ($F_{1,10} = 10.79$, $P = 0.006$; Fig. 4a) vesicle release, without preventing spiking at the soma (Supplementary Fig. 14). BLA somata inhibition did not induce an anxiogenic response, perhaps owing to the simultaneous decrease in direct BLA–CeM excitatory input. We also found that the eNpHR3.0:BLA–CeA group showed reduced open-arm time and probability of open-arm entry on the EPM ($F_{1,40} = 21.08$, $P < 0.00001$; and $F_{1,40} = 19.93$, $P < 0.00001$, respectively; Fig. 4b, c) and centre time in the OFT ($F_{1,100} = 18.919$, $P < 0.00001$; Fig. 4d, e) during photostimulation when compared to the eYFP and BLA(somata) groups, without altering locomotor activity (Supplementary Fig. 15). These data demonstrate that preferential inhibition of BLA–CeL synapses acutely increases anxiety-like behaviours.

Here, we have identified the BLA–CeL pathway as a neural substrate for real-time bidirectional modulation of the unconditioned expression of anxiety. The observation that selective illumination of specific BLA terminals produces distinct, and even opposite, behavioural responses

To confirm further that the anxiolytic effect was due to the selective activation of BLA–CeL projections alone, and not BLA axons passing through the CeA or back-propagation of action potentials to BLA cell bodies that would then innervate all BLA projection target regions, we tested whether local glutamate receptor antagonism would attenuate light-induced anxiolytic effects. This distinction is critical, as there have been previous reports that CeA lesions that alter anxiety are confounded by the ablation of BLA projections to the BNST passing through the CeA²⁸. In a separate group of mice, we selectively illuminated BLA–CeA terminals as before ($n = 8$; Supplementary Fig. 1), but infused glutamate antagonists or saline via the fibre-optic guide cannula before testing on the EPM and OFT. Confirming a local synaptic mechanism rather than control of fibres of passage, intra-CeA glutamate receptor antagonism abolished light-induced reductions in anxiety as measured by open-arm time ($F_{1,35} = 8.61$, $P = 0.008$) and probability of open-arm entry on the EPM ($F_{1,35} = 5.92$, $P = 0.02$), and centre time during the OFT ($F_{1,77} = 13.99$, $P = 0.0006$; Fig. 3c, d). Importantly,

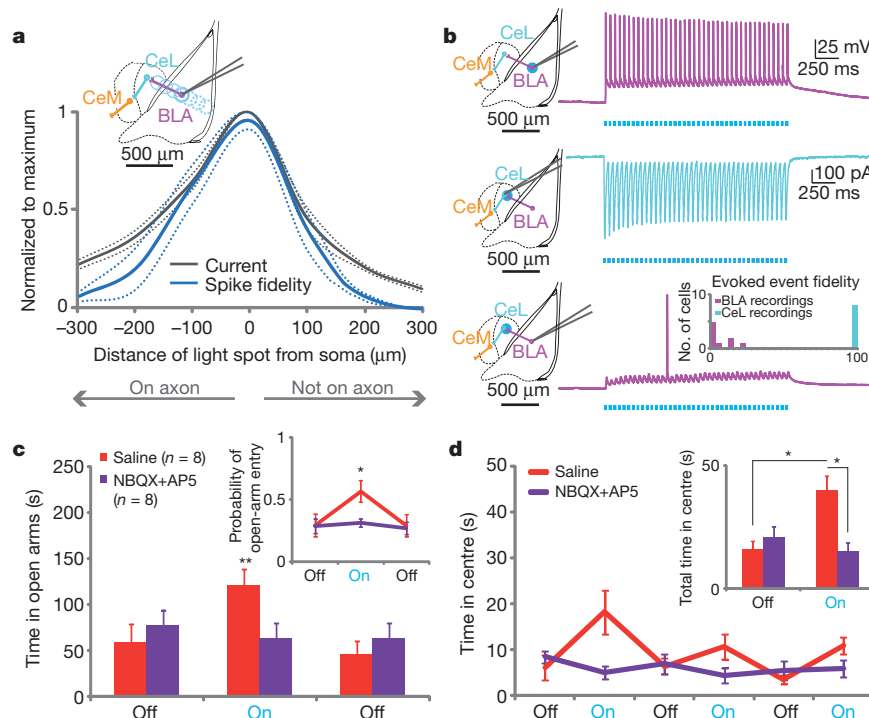


Figure 3 | Light-induced anxiolytic effects are attributable to activation of BLA–CeA synapses. **a**, **b**, Schematic of the recording site and illumination positions as whole-cell recordings were performed at each illumination location in 100- μ m increments away from the cell soma both over a visualized axon and in a direction that was not over an axon (inset). Normalized summary of spike fidelity and depolarizing current (**a**) to a 20-Hz train delivered at various distances from the soma. **b**, Representative traces on ~ 125 - μ m-diameter illumination at various locations within each slice ($n = 7$). Illumination of BLA somata elicits high-fidelity spiking (top). Illumination of BLA terminals in CeL elicits strong excitatory responses shown in voltage-clamp in the postsynaptic

CeL neuron (middle), but does not elicit reliable antidromic spiking in the BLA neuron itself (bottom), summarized in a frequency histogram (inset, 120 pulses per cell). **c**, **d**, A separate group of ChR2:BLA–CeA mice ($n = 8$) performed the EPM and OFT twice, one session preceded with intra-CeA infusions of saline (red) and the other session with glutamate receptor antagonists NBQX and AP5 (purple), counter-balanced for order. Glutamate receptor blockade in the CeA attenuated light-induced increases in both open-arm time (**c**) and probability of open-arm entry (inset) on the EPM and centre time on the OFT (**d**, inset shows pooled summary), without altering baseline performance.

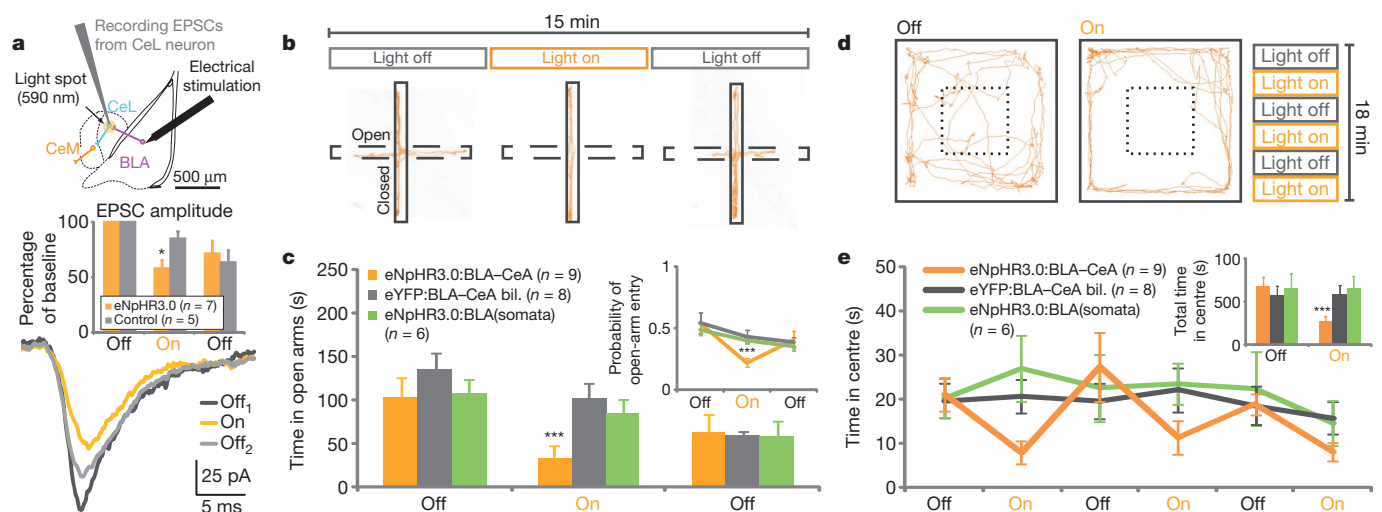


Figure 4 | Selective inhibition of BLA terminals in the CeA induces an acute and reversible increase in anxiety. **a–e**, Mice were group-housed in a low-stress environment and received bilateral constant 594-nm light during light-on epochs. **a**, Selective illumination of eNpHR3.0-expressing BLA terminals suppresses vesicle release evoked by electrical stimulation in the BLA. Schematic indicates the locations of the stimulating electrode, the recording electrode and the ~ 125 - μ m diameter light spot. Representative CeL EPSCs before (Off₁), during (On) and after (Off₂) selective illumination of eNpHR3.0-expressing BLA terminals. Normalized EPSC amplitude summary data from sections containing BLA neurons expressing eNpHR3.0 ($n = 7$) and non-

transduced controls ($n = 5$) show that selectively illuminating BLA–CeL terminals reduces ($*P = 0.006$) electrically evoked EPSC amplitude in postsynaptic CeL neurons relative to non-transduced control slice preparations (inset). **b**, **c**, Representative eNpHR3.0:BLA–CeA path (**b**) indicates reduced open-arm time (**c**) and probability of open arm entry (inset) during illumination, relative to controls. bil., bilateral. **d**, **e**, Representative eNpHR3.0:BLA–CeA path (**d**) reflects reduced centre time on the OFT (**e**) for the eNpHR3.0:BLA–CeA group during light-on, but not light-off, epochs as compared to controls (inset shows pooled data).

from illumination of all glutamatergic BLA somata nonspecifically, points to the essential value of optogenetic control in causally dissecting intact neural circuitry, and indicates that multiple subpopulations or projections of BLA neurons can act in opposition (for example, direct excitation of CeM along with feed-forward inhibition of CeM). Neural circuitry arranged in this way provides many opportunities for modulation of expression of anxiety phenotypes; for example, this microcircuit is well-positioned to be influenced by top-down cortical control from regions important for processing fear and anxiety, including the prelimbic, infralimbic, anterior cingulate and insular cortices that provide robust innervation to the BLA and CeL.

These data are consistent with reports implicating CeA involvement in anxiety^{9,11,12}, but it is important to note that our findings do not exclude downstream or parallel circuits including the BNST²⁸, the insular and prefrontal cortices²⁹, and the septal-hippocampal circuit³⁰; for example, stress induces CeL release of corticotropin releasing hormone (CRH) in the BNST²⁸. In the course of providing insight into native anxiogenic and anxiolytic processes, these findings demonstrate that anxiety is continuously regulated by balanced antagonistic pathways within the amygdala, and illustrate the importance of resolving specific projections in the study of neural circuit function relevant to psychiatric disease.

METHODS SUMMARY

Virus-mediated opsin gene expression. The pAAV-CaMKII α -hChR2(H134R)-eYFP, pAAV-CaMKII α -eYFP and pAAV-CaMKII α -eNpHR3.0-eYFP plasmids were designed and constructed by standard methods and packaged as AAV₅. Virus (0.5 μ l) was injected into the BLA. Maps and clones are available at <http://www.optogenetics.org>.

In vivo projection-specific targeting. To investigate the role of the BLA–CeL pathway in modulating anxiety, we performed viral transduction and surgical implantation of bevelled guide cannulae to allow selective illumination of BLA fibres in the CeA under stereotaxic guidance. Behavioural, electrophysiological and imaging data were collected 4–6 weeks after surgery.

Two-photon imaging and functional mapping using ex vivo electrophysiology. Acute slices were collected for two-photon imaging and ex vivo electrophysiological recordings. While light-stimulation parameters used *in vivo* were delivered via fibre optics, and light in ex vivo experiments was delivered onto coronal sections, we matched light power density at our target region ~ 6 mW mm⁻². Whole-cell recordings were made from BLA pyramidal neurons simultaneously during two-photon visualization of neuronal processes with Alexa Fluor dye. We visually tracked axonal projections from BLA neurons to the CeL nucleus. We recorded from CeL neurons on illumination with an aperture-restricted light spot (~ 125 μ m diameter), mimicking the preferential illumination of BLA terminals, but not BLA somata, delivered *in vivo*. Two-photon imaging allowed axonal tracking to the CeM, where whole-cell recordings were collected from CeM neurons in the terminal field of CeL axons, with aperture-restricted illumination over the CeL to allow selective illumination of BLA terminals in the CeL while recording from the CeM neuron.

Opsin expression validation and immunohistochemistry. To validate specificity, sensitivity and spatial distribution of opsin expression as well as neuronal activity, brain slices were prepared for optical microscopy and immunohistochemistry. Coronal sections were stained for 4',6-diamidino-2-phenylindole (DAPI) and immunoreactivity for *c-fos*. Quantitative analyses of confocal images were performed with both staining and analysis blind to experimental condition.

Received 7 November 2010; accepted 14 January 2011.

Published online 9 March 2011.

1. Lieb, R. Anxiety disorders: clinical presentation and epidemiology. *Handb. Exp. Pharmacol.* **169**, 405–432 (2005).
2. Kessler, R. C. *et al.* Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593–602 (2005).
3. Koob, G. F. Brain stress systems in the amygdala and addiction. *Brain Res.* **1293**, 61–75 (2009).
4. Ressler, K. J. & Mayberg, H. S. Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. *Nature Neurosci.* **10**, 1116–1124 (2007).
5. LeDoux, J. The emotional brain, fear, and the amygdala. *Cell. Mol. Neurobiol.* **23**, 727–738 (2003).
6. Pare, D., Quirk, G. J. & LeDoux, J. E. New vistas on amygdala networks in conditioned fear. *J. Neurophysiol.* **92**, 1–9 (2004).

7. Tye, K. M., Stuber, G. D., de Ridder, B., Bonci, A. & Janak, P. H. Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. *Nature* **453**, 1253–1257 (2008).
8. Weiskrantz, L. Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *J. Comp. Physiol. Psychol.* **49**, 381–391 (1956).
9. Kalin, N. H., Shelton, S. E. & Davidson, R. J. The role of the central nucleus of the amygdala in mediating fear and anxiety in the primate. *J. Neurosci.* **24**, 5506–5515 (2004).
10. Lesscher, H. M. *et al.* Amygdala protein kinase C epsilon regulates corticotropin-releasing factor and anxiety-like behavior. *Genes Brain Behav.* **7**, 323–333 (2008).
11. Etkin, A., Prater, K. E., Schatzberg, A. F., Menon, V. & Greicius, M. D. Disrupted amygdalar subregion functional connectivity and evidence of a compensatory network in generalized anxiety disorder. *Arch. Gen. Psychiatry* **66**, 1361–1372 (2009).
12. Lyons, A. M. & Thiele, T. E. Neuropeptide Y conjugated to saporin alters anxiety-like behavior when injected into the central nucleus of the amygdala or basomedial hypothalamus in BALB/cJ mice. *Peptides* **31**, 2193–2199 (2010).
13. Roozendaal, B., McEwen, B. S. & Chattarji, S. Stress, memory and the amygdala. *Nature Rev. Neurosci.* **10**, 423–433 (2009).
14. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8**, 1263–1268 (2005).
15. Gradinaru, V. *et al.* Molecular and cellular approaches for diversifying and extending optogenetics. *Cell* **141**, 154–165 (2010).
16. Deisseroth, K. Optogenetics: controlling the brain with light. *Sci. Am.* **303**, 48–55 (2010).
17. Woods, J. H., Katz, J. L. & Winger, G. Benzodiazepines: use, abuse, and consequences. *Pharmacol. Rev.* **44**, 151–347 (1992).
18. Ciochi, S. *et al.* Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature* **468**, 277–282 (2010).
19. Haubensack, W. *et al.* Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature* **468**, 270–276 (2010).
20. Carlsen, J. Immunocytochemical localization of glutamate decarboxylase in the rat basolateral amygdaloid nucleus, with special reference to GABAergic innervation of the amygdaloatrial projection neurons. *J. Comp. Neurol.* **273**, 513–526 (1988).
21. Smith, Y. & Pare, D. Intra-amygdaloid projections of the lateral nucleus in the cat: PHA-L anterograde labeling combined with postembedding GABA and glutamate immunocytochemistry. *J. Comp. Neurol.* **342**, 232–248 (1994).
22. McDonald, A. J. Cytoarchitecture of the central amygdaloid nucleus of the rat. *J. Comp. Neurol.* **208**, 401–418 (1982).
23. Krettek, J. E. & Price, J. L. A description of the amygdaloid complex in the rat and cat with observations on intra-amygdaloid axonal connections. *J. Comp. Neurol.* **178**, 255–279 (1978).
24. Krettek, J. E. & Price, J. L. Amygdaloid projections to subcortical structures within the basal forebrain and brainstem in the rat and cat. *J. Comp. Neurol.* **178**, 225–253 (1978).
25. Davis, M. *in The Amygdala: A Functional Analysis* (ed. Aggleton, J. P.) 213–288 (Oxford Univ. Press, 2000).
26. Pitkanen, A. *in The Amygdala: A Functional Analysis* (ed. Aggleton, J. P.) 31–99 (Oxford Univ. Press, 2000).
27. Carola, V., D'Olimpio, F., Brunamonti, E., Mangia, F. & Renzi, P. Evaluation of the elevated plus-maze and open-field tests for the assessment of anxiety-related behaviour in inbred mice. *Behav. Brain Res.* **134**, 49–57 (2002).
28. Davis, M., Walker, D. L., Miles, L. & Grillon, C. Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology* **35**, 105–135 (2010).
29. Shin, L. M. & Liberzon, I. The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology* **35**, 169–191 (2010).
30. Gray, J. A. & McNaughton, N. The neuropsychology of anxiety: reprise. *Nebr. Symp. Motiv.* **43**, 61–134 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank P. Janak, H. Fields, G. Stuber, E. Thomas, F. Zhang, I. Witten, V. Sohal, T. Davidson and M. Warden as well as J. Mattis, R. Durand, M. Mogri, J. Mirzabekov and E. Steinberg for discussions, and the entire K.D. laboratory for their support. All viruses were packaged at University of North Carolina (UNC) Vector Core. Supported by NIMH (1F32MH088010-01, K.M.T.), NARSAD (K.R.T.), Samsung Scholarship (S.-Y.K.), NSF IGERT Award 0801700 (L.G.) and the Defense Advanced Research Projects Agency Reorganization and Plasticity to Accelerate Injury Recovery (N66001-10-C-2010), the Alice Woo, Albert Yu, Snyder, and McKnight Foundations, as well as NIDA, NIMH and the NIH Pioneer Award (K.D.).

Author Contributions K.M.T., R.P., S.-Y.K., L.E.F. and K.D. contributed to study design and data interpretation. K.M.T., R.P., S.-Y.K. and L.E.F. contributed to data collection and K.M.T. coordinated data collection and analysis. K.M.T., S.-Y.K., H.Z. and K.R.T. contributed to immunohistochemical processing, fluorescence imaging and quantitative analyses. K.M.T. and L.G. performed the behavioural and ex vivo electrophysiology statistical analyses. V.G. and C.R. contributed to the design of eNpHR3.0. C.R. cloned all constructs and managed viral packaging processes. K.D. supervised all aspects of the work. All authors contributed to writing the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.D. (deissero@stanford.edu).

Catalytic activity of the caspase-8-FLIP_L complex inhibits RIPK3-dependent necrosis

Andrew Oberst¹, Christopher P. Dillon¹, Ricardo Weinlich¹, Laura L. McCormick¹, Patrick Fitzgerald¹, Cristina Pop², Razq Hakem³, Guy S. Salvesen² & Douglas R. Green¹

Caspase-8 has two opposing biological functions—it promotes cell death by triggering the extrinsic pathway of apoptosis, but also has a survival activity, as it is required for embryonic development¹, T-lymphocyte activation², and resistance to necrosis induced by tumour necrosis factor- α (TNF- α) and related family ligands^{3,4}. Here we show that development of caspase-8-deficient mice is completely rescued by ablation of receptor interacting protein kinase-3 (RIPK3). Adult animals lacking both caspase-8 and RIPK3 display a progressive lymphoaccumulative disease resembling that seen with defects in CD95 or CD95-ligand (also known as FAS and FASLG, respectively), and resist the lethal effects of CD95 ligation *in vivo*. We have found that caspase-8 prevents RIPK3-dependent necrosis without inducing apoptosis by functioning in a proteolytically active complex with FLICE-like inhibitory protein long (FLIP_L, also known as CFLAR), and this complex is required for the protective function.

The death receptor pathway of apoptosis is induced by ligation of a subset of tumour necrosis factor (TNF) receptor super-family members (the death receptors)⁵. This pathway involves the recruitment of an adaptor molecule, FADD, to the intracellular region of the receptor; FADD in turn binds and thereby activates caspase-8 to initiate apoptosis. Cell death in this pathway is antagonized by another protein, FLIP_L (herein called FLIP), which resembles caspase-8 but lacks a catalytic site⁵.

Intriguingly, genetic ablation of caspase-8 (ref. 1), FADD⁶ or FLIP⁷ results in embryonic lethality around embryonic day E10.5, showing that these proteins function in cell survival as well as cell death. This is supported by the finding that caspase-8 deficiency by short interfering RNA (siRNA) knockdown or gene ablation sensitizes fibroblasts for necrotic cell death in response to TNF³ (Supplementary Fig. 1a). Necrosis induced by TNF in the presence of caspase inhibitors is dependent on the kinase activity of receptor-interacting protein kinase-1 (RIPK1)^{4,8} and RIPK3 (refs 9–11), although the mechanisms remain obscure.

To determine if RIPK3-dependent necrosis contributes to the embryonic lethality of caspase-8-deficient mice, we generated *Casp8*^{-/-}*Ripk3*^{-/-} double knockout (DKO) animals. Although, as anticipated, we were unable to obtain viable *Casp8*^{-/-}*Ripk3*^{WT} mice, *Casp8*^{-/-}*Ripk3*^{-/-} DKO mice were born at the expected frequency (Fig. 1a). These animals displayed no gross developmental abnormalities (Supplementary Fig. 1b), and their mass at different ages was indistinguishable from that of heterozygous mice (Supplementary Fig. 1c), as described for the *Ripk3*^{-/-} mouse¹², despite lacking detectable caspase-8 or RIPK3 (Supplementary Fig. 1d). Thymocytes from these animals underwent apoptosis in response to several agents known to induce the mitochondrial pathway of apoptosis, but were resistant to apoptosis induced by ligation of the death receptor CD95 (Supplementary Fig. 1e). We examined the latter effect in more detail, as injection of agonistic anti-CD95 antibody is known to trigger hepatocyte apoptosis, liver damage, and death in wild-type mice¹³. Whereas anti-CD95 caused liver

destruction and mortality in heterozygous *Casp8*^{+/-}*Ripk3*^{-/-} animals, *Casp8*^{-/-}*Ripk3*^{-/-} mice were completely resistant to this insult (Fig. 1b and Supplementary Fig. 1f–h).

In young *Casp8*^{-/-}*Ripk3*^{-/-} DKO mice, lymphoid organs appeared overtly normal (Supplementary Fig. 2a), T-lymphocyte proliferation in response to activation was identical to that of heterozygote littermates (Supplementary Fig. 2b), and T cells from these animals displayed expansion and subsequent peripheral deletion *in vivo* when challenged with the bacterial superantigen *Staphylococcus* enterotoxin B (SEB) (Supplementary Fig. 2c). However, we noted that older DKO mice displayed a striking lymphoaccumulation (Fig. 2a), resembling that described in animals lacking functional CD95 or CD95-ligand¹⁴. The latter are known to accumulate an unusual population of B220⁺, CD3⁺, CD4⁻, CD8⁻ T lymphocytes, also seen in humans with defective CD95 or CD95-ligand¹⁴. Whereas young (1 month) *Casp8*^{-/-}*Ripk3*^{-/-} DKO mice showed normal mature T lymphocyte subsets, we observed a marked increase in B220⁺, CD3⁺ cells as the animals aged (Fig. 2b, Supplementary Fig. 2d).

The ability of RIPK3 ablation to rescue the lethal phenotype of caspase-8 deletion strongly indicates that caspase-8-mediated inhibition of RIPK3-dependent necrosis is necessary for embryonic development, and that this is the primary protective role of caspase-8 in development. This raises the question of how caspase-8 can mediate this effect without itself engaging apoptotic cell death in the cells in

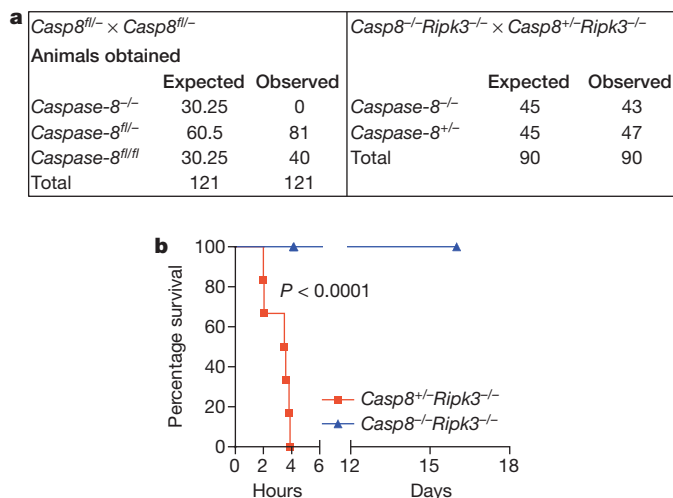


Figure 1 | *Casp8*^{-/-}*Ripk3*^{-/-} mice are viable and functionally deficient for caspase-8. **a**, Expected and observed frequency of caspase-8 status in offspring from crosses of mice with the indicated genotypes. 'fl' indicates an allele of caspase-8 that is present but flanked with LoxP sites. ($P < 0.0001$ left, $P = 0.6733$ right). **b**, Effect of tail vein injection of 15 μ g per animal of the CD95-activating antibody Jo2 on mice of the indicated genotypes. ($n = 6$ *Casp8*^{+/-}*Ripk3*^{-/-}, 8 DKO).

¹Dept. of Immunology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ²Program in Apoptosis and Cell Death Research, Sanford-Burnham Medical Research Institute, La Jolla, California 92037, USA. ³Ontario Cancer Institute, University of Toronto, Toronto, Ontario M5G 2M9, Canada.

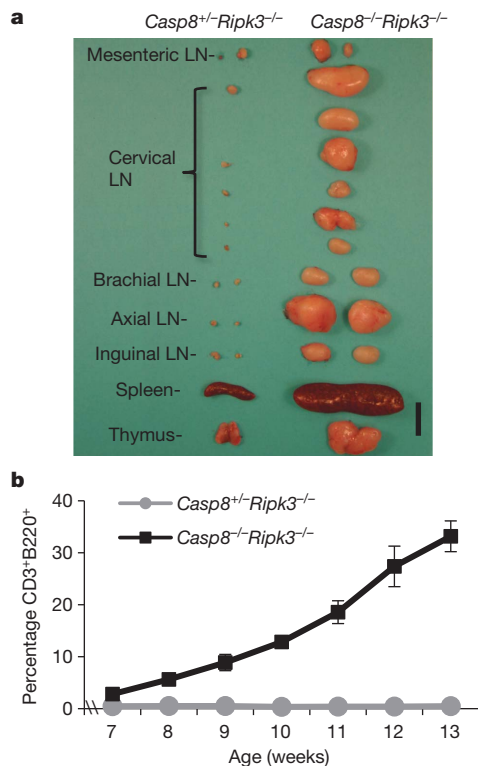


Figure 2 | *Casp8*^{-/-}*Ripk3*^{-/-} mice display progressive severe lymphoaccumulation. **a**, Lymphoid organs removed from 15-week-old littermate mice of the indicated genotypes. LN, lymph node. Scale bar, 1 cm. **b**, Percentage of total blood cells (following red blood cell lysis) that are B220⁺CD3⁺ in mice of the indicated genotypes and ages. Error bars, s.d.; *n* = 3 each genotype. Both **a** and **b** are representative of similar results obtained from all sampled mice of the indicated genotypes.

which it manifests this protective function. A clue is provided by the observation that a mutant of caspase-8, lacking the cleavage site between the large and small subunits of the mature enzyme, rescued survival of caspase-8-deficient animals when expressed as a bacterial artificial chromosome (BAC)-transgene¹⁵. Such 'non-cleavable' caspase-8 has been shown to be unable to restore death-receptor-induced apoptosis in caspase-8-deficient cells^{16,17}. Biochemical and structural studies have indicated that FLIP can heterodimerize with caspase-8 in kosmotropic salt, and that this complex may be able to impart catalytic activity on caspase-8 in the absence of interdomain cleavage^{18,19}. We sought to test this directly by examining enzymatic activity of non-cleavable caspase-8^{DA} when dimerized with FLIP *in vitro*. We generated caspase-8 or non-cleavable caspase-8^{DA}, as well as FLIP, in which the prodomains were replaced by FKBP or FRB domains, respectively, allowing enforced homo- or heterodimerization upon addition of specific FK506 derivatives²⁰. Although homodimerized FKBP-caspase-8^{DA} was catalytically inactive under physiological buffer conditions (Supplementary Fig. 3a)¹⁶, heterodimerization of this non-cleavable caspase-8 with FRB-FLIP was enzymatically active (Fig. 3a). Thus, FLIP can impart catalytic activity to non-cleavable caspase-8.

Because ablation of FLIP leads to developmental defects similar to those observed upon caspase-8 ablation⁷, and because expression of non-cleavable caspase-8^{DA} allows normal development, we proposed that FLIP may activate caspase-8 to allow suppression of RIPK3-dependent death. To investigate this possibility, we generated mouse embryonic fibroblasts (MEF) from *Casp8*^{-/-}*Ripk3*^{-/-} DKO embryos and reconstituted them with caspase-8, non-cleavable caspase-8^{DA} or catalytically inactive caspase-8^{CA}, plus or minus RIPK3 (Fig. 3b). We then knocked down FLIP expression by siRNA (Supplementary Fig. 3b). Cells lacking RIPK3 underwent cell death upon treatment with TNF only if they expressed wild-type caspase-8, and only if FLIP was lacking

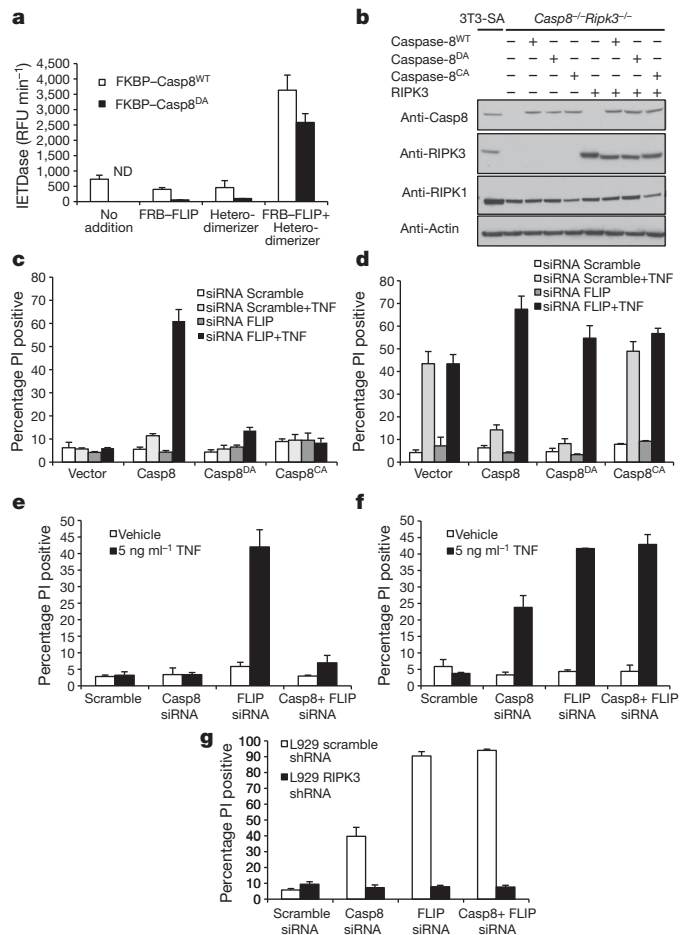


Figure 3 | Both catalytically active caspase-8 and FLIP are required for suppression of TNF-induced RIPK3-dependent death. **a**, Fluorogenic substrate cleavage activity of recombinant purified FKBP-caspase-8^{WT} or non-cleavable FKBP-caspase-8^{DA} in the presence of recombinant purified FRB-FLIP, a compound that induces FKBP-FRB heterodimers (heterodimerizer), or FRB-FLIP and heterodimerizer. ND, none detected. Error bars, s.d.; *n* = 3. **b**, Western blot analysis of *Casp8*^{-/-}*Ripk3*^{-/-} (DKO) MEF stably expressing the indicated mutants of caspase-8 and RIPK3. Caspase-8^{CA} indicates catalytically inactive caspase-8. **c**, **d**, Cell death assessed by propidium iodide (PI) uptake of DKO MEF expressing the indicated caspase-8 mutants in the absence (**c**) or presence (**d**) of stably expressed RIPK3, following transfection with scramble or FLIP-targeted siRNA, and 12 h TNF treatment. Error bars, s.d.; *n* = 3. **e**, **f**, Cell death (PI uptake) of SVEC 4-10 cells stably expressing RIPK3-specific (**e**) or scramble (**f**) shRNA, transfected with the indicated siRNAs and treated with TNF for 12 h. Graph is mean of two independent experiments, error bars indicate range. **g**, Cell death (PI uptake) of L929 cells expressing scramble or RIPK3-specific shRNA, transfected with siRNAs specific for caspase-8 and/or FLIP as indicated. Death was assessed 48 h post-transfection. Error bars, s.d.; *n* = 3. The data presented are representative of results obtained with either of two separate siRNAs to both caspase-8 and FLIP.

(Fig. 3c). Conversely, cells expressing RIPK3 died upon TNF exposure if they lacked caspase-8 or expressed catalytically inactive caspase-8^{CA} (Fig. 3d). Importantly, cells expressing non-cleavable caspase-8^{DA} were resistant to TNF-induced, RIPK3-dependent death, but became sensitive to this form of cell death upon knockdown of FLIP (Fig. 3d). Thus, FLIP expression prevents caspase-8-dependent, RIPK3-independent death (apoptosis), and prevents RIPK3-dependent, necrotic cell death in a caspase-8-dependent manner. Apoptosis mediated by caspase-8 depends on cleavage of the caspase^{16,17}, whereas inhibition of RIPK3-mediated cell death does not.

We obtained similar results in several cell lines. In murine SVEC cells, knockdown of caspase-8 or FLIP promoted TNF-induced cell

death, which in the case of caspase-8 was completely blocked by concomitant knockdown of RIPK3 (Fig. 3e, f and Supplementary Fig. 3c). In contrast, TNF-induced cell death promoted by knockdown of FLIP was only partially dependent on RIPK3, presumably due to compensation by caspase-8-mediated apoptosis. This was supported by the combined knockdown of caspase-8 and FLIP, in which TNF-induced cell death was entirely RIPK3 dependent (Fig. 3e, f). Similar results were obtained in two 3T3 cell lines with (SA) or without (NIH) RIPK3 (Supplementary Fig. 3d–f).

In L929 cells, treatment with caspase inhibitor alone has been shown to cause RIPK3-dependent death due to autocrine production of TNF¹⁶. In these cells, knockdown of either caspase-8 or FLIP led to cell death that was fully dependent on RIPK3 (Fig. 3g and Supplementary Fig. 3g). For unknown reasons, no RIPK3-independent, caspase-8-dependent apoptosis was observed in these cells. However, together with the results obtained in *Casp8*^{-/-}*Ripk3*^{-/-} DKO MEF expressing non-cleavable caspase-8^{DA} (discussed above), these observations indicate that rather than simply functioning to dampen caspase-8 activation, FLIP is an important component of the protective effect.

Crma is a pox-virus protein that potently inhibits the enzymatic activity of caspase-8 homodimers²¹ (Fig. 4a). Remarkably, the ability of Crma to inhibit the activity of caspase-8^{DA}–FLIP heterodimers was strikingly attenuated (Fig. 4a). Proposing that this property would allow discrimination between the cellular roles of caspase-8 homodimers and caspase-8–FLIP heterodimers, we expressed Crma in RIPK3-deficient 3T3-NIH cells or RIPK3-expressing 3T3-SA cells and examined the response to TNF (Fig. 4b). Crma blocked the RIPK3-independent, caspase-8-dependent apoptotic death induced by TNF

treatment of 3T3-NIH cells following FLIP knockdown. However, Crma did not sensitize RIPK3-expressing 3T3-SA cells to TNF-induced death (Fig. 4b), demonstrating that the inhibitor is not able to influence the ‘survival’ role of caspase-8. However, 3T3-SA cells expressing Crma were readily sensitized to undergo TNF-dependent death by knockdown of caspase-8, FLIP, or both (Fig. 4d). Importantly, death induced by FLIP knockdown in 3T3-SA cells was only partially inhibited by the RIPK1 inhibitor necrostatin-1 (Nec1) or by stable knockdown of RIPK3, presumably due to contribution from caspase-8-dependent, RIPK3-independent apoptotic death under these conditions (Fig. 4c and Supplementary Fig. 4a). However, in Crma-expressing 3T3-SA cells, the death induced by FLIP knockdown and TNF treatment was entirely inhibitable by Nec1 (Fig. 4d), presumably due to Crma-mediated blockade of caspase-8-dependent apoptosis in these cells. These observations provide further support for the idea that the caspase-8–FLIP heterodimer is functionally active in inhibiting RIPK3-dependent necrosis without promoting apoptosis.

We next sought to determine whether FLIP itself is required for the suppression of RIPK3-dependent TNF induced death, or whether it is merely required to dampen the apoptotic effect of caspase-8. To differentiate between these possibilities, we blocked the mitochondrial pathway of apoptosis in 3T3-SA cells by expression of the anti-apoptotic Bcl-2 family member Bcl-XL (also known as Bcl2l1), which blocks caspase-8-mediated apoptosis in many cells²². Strikingly, knockdown of FLIP in these cells strongly sensitized them to TNF-induced cell death, an effect that was completely inhibited by Nec1 or concomitant RIPK3 knockdown (Fig. 4e and Supplementary Fig. 4b). Western blot analysis showed appearance of fully-processed caspase-8 upon FLIP knockdown and TNF treatment in 3T3-SA cells with or without over-expression of Bcl-XL (Supplementary Fig. 4c). However, Bcl-XL expression blocked appearance of fully processed caspase-3 under these conditions, consistent with inhibition of the mitochondrial pathway of apoptosis (Supplementary Fig. 4c). Therefore, cells in which caspase-8 is activated but apoptosis is blocked nevertheless require FLIP for effective suppression of RIPK3-mediated death. Immunoprecipitation of FADD following TNF treatment of Bcl-XL expressing 3T3 cells did not co-precipitate RIPK1 or RIPK3 (Fig. 4f). However, when caspase-8, FLIP, or both were knocked down in these cells, association of both RIPK1 and RIPK3 with FADD was observed (Fig. 4f). Cleaved caspase-8 was observed in whole lysates from these cells upon FLIP knockdown and TNF treatment, indicating that active caspase-8 was present in these cells, but was unable to prevent formation of the RIPK1–RIPK3 complex in the absence of FLIP (Supplementary Fig. 4d). Therefore, caspase-8–FLIP heterodimer, but not caspase-8 homodimer, prevents the stable association of FADD, RIPK1 and RIPK3, and prevents necrotic death.

The death of mice homozygous for caspase-8 deletion at E10.5 has been traced to a failure in early vascularization and haematopoietic development¹. Our results demonstrate that this requisite role for caspase-8 in embryogenesis is eliminated by concurrent ablation of RIPK3. As RIPK3 is essential for programmed necrosis, the likely role of caspase-8 in development is to suppress the lethal effects of RIPK3, probably associated with necrotic death of endothelial and haematopoietic precursors. Consistent with this idea, we observed RIPK3 expression in haematopoietic tissues (Supplementary Fig. 1d). Further, we found that the ability of caspase-8 to inhibit RIPK3-dependent necrosis depends on the expression of FLIP. Caspase-8–FLIP heterodimers are enzymatically active, and this does not require cleavage of caspase-8 between its large and small subunits. In contrast, apoptosis induced by caspase-8 depends on such cleavage, which stabilizes the homodimer^{15–17,23}. This explains the ability of non-cleavable caspase-8 to rescue development in caspase-8-deficient mice without restoring sensitivity to death receptor-induced apoptosis¹⁵. This also explains the observation that activation of caspase-8-deficient T-cells causes cell death that is inhibitable by Nec1²⁴. Our finding that T cells lacking both caspase-8 and RIPK3 proliferate normally is consistent with the idea that caspase-8 functions to suppress RIPK1–RIPK3-dependent

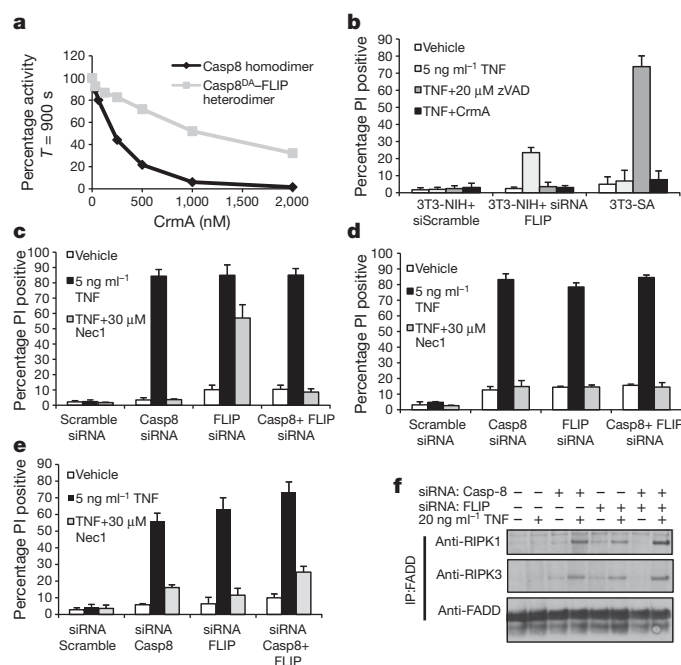


Figure 4 | The caspase-8–FLIP heterodimeric complex suppresses RIPK3-dependent cell death. **a**, Inhibitory effect of CrmA on FKBP–caspase-8 homodimers or caspase-8^{DA}–FLIP heterodimers induced by dimerizer treatment of purified recombinant protein. **b**, Cell death (PI uptake) of RIPK3-deficient (NIH) or RIPK3-expressing (SA) 3T3 cells treated as indicated for 12 h. Error bars, s.d.; *n* = 3. **c–e**, Cell death (PI uptake) of 3T3-SA cells stably expressing vector (**c**) or CrmA (**d**), or anti-apoptotic Bcl-XL (**e**) following transfection with the indicated siRNA and treatment with TNF and the RIPK1 inhibitor Nec1 as indicated for 24 h. Error bars, s.d.; *n* = 3. **f**, 3T3-SA cells expressing Bcl-XL were subjected to immunoprecipitation (IP) of FADD following transfection of siRNAs to caspase-8 and FLIP and treatment with TNF for 90 min as indicated. Immune complexes were resolved by western blotting with the indicated antibodies. The data presented are representative of results obtained with either of two separate siRNAs to both caspase-8 and FLIP.

cell death. Demonstration that non-cleavable caspase-8 can rescue T-cell proliferation¹⁵, along with our finding that FLIP can impart catalytic activity to this form of caspase-8, further supports an essential role for the caspase-8–FLIP complex in this process. Furthermore, because FLIP is a well known transcriptional target of NF- κ B²⁵, the suppression of RIPK3-dependent necrosis may represent a novel role for NF- κ B signalling in immune cell proliferation. Taken together, these results support the idea that the main non-apoptotic function of caspase-8 is to suppress RIPK3-dependent necrosis during development and immune cell proliferation, and that it does so in complex with FLIP. Although caspase-8 may yet prove to have roles in other cellular processes, such as NF- κ B activation²⁶ and cell motility²⁷, our results do not provide support for such activities in development.

On the basis of the phenotypes of the caspase-8 (ref. 1), FADD⁶, and FLIP⁷-deficient animals, it is likely that the complex that protects cells from RIPK3-dependent necrosis includes all three of these components. Such complexes readily form in response to death receptor signalling²⁸. At present, it is not known if development of caspase-8 deficient animals can be rescued by deletion of TNFR1, CD95, or the murine TRAIL receptor, all of which are capable of triggering RIPK-dependent necrosis in the absence of caspase-8 (ref. 4). It is also possible that these (and their ligands) function redundantly to cause necrosis in caspase-8-deficient endothelium and haematopoietic cells, or that other death or immune receptors cause this effect.

The ability of CrmA to block caspase-8-mediated apoptosis but not the protective effect of caspase-8–FLIP indicates that some viruses can subvert this system to prevent both apoptosis and necrosis of infected cells. Consistent with our findings, T cells expressing a CrmA transgene were observed to resist death receptor-induced apoptosis, but unlike caspase-8-deficient T cells, showed no proliferative defects in response to activation²⁹. In contrast, in one study CrmA overexpression was observed to sensitize L929 cells to TNF-induced necrosis; it is likely that in this case CrmA was expressed at high enough levels that its relatively weak inhibition of caspase-8–FLIP activity was sufficient to disrupt protection³.

The precise mechanism by which the catalytic activity of the caspase-8–FLIP complex is engaged to prevent RIPK3-dependent necrosis without triggering apoptosis is not presently known. Our data indicate that both caspase-8 and FLIP are required to disrupt formation of a stable complex containing FADD, RIPK1 and RIPK3 following TNF ligation. In cell-free systems, caspase-8–FLIP heterodimers display less proteolytic activity than caspase-8 homodimers on apoptotic substrates such as Bid and caspase-3 (refs 17, 30), while maintaining a low but perceptible ability to cleave RIPK1 (ref. 30). It is therefore possible that FADD-dependent formation of caspase-8–FLIP heterodimers prevents stable RIPK1–RIPK3 association by cleaving RIPK1. However, the target and efficiency of this protease *in vivo* is likely to be determined by its recruitment to specific complexes, and the presence and availability of substrates therein, making it difficult to draw firm conclusions from biochemical studies. It remains possible that downstream targets of RIPK3 that cause necrosis (currently unknown) are the relevant substrates. NADPH-oxidase, mitochondrial reactive oxygen species production³, and metabolic enzymes⁹ have been suggested to be possible downstream effectors of RIPK3, but at present, we do not know how these may be engaged by this kinase following its activation by RIPK1.

METHODS SUMMARY

Mice and treatments. Mice with a deleted allele of caspase-8 were generated by germ line deletion of a caspase-8^{fllox} allele described previously². RIPK3-deficient animals were obtained from V. Dixit¹². Genotypes were confirmed by tail snip PCR as described previously. For Jo2 injections, animals were injected via tail vein with 15 μ g purified Jo2 in lipopolysaccharide-free PBS per animal. Liver enzymes were assayed using a Trilogy Multi-Purpose Analyzer System from Drew Scientific, and liver sections were created and stained with haematoxylin and eosin, in the St. Jude Veterinary Pathology Core facility. For SEB injections, 50 μ g SEB (Toxin Technology Inc.) per animal was injected via tail vein and T-cell populations were monitored by retro-orbital bleed and FACS as detailed

previously. The St. Jude Institutional Animal Care and Use Committee approved all procedures in accordance with the Guide for the Care and Use of Animals.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 November 2010; accepted 21 January 2011.

Published online 2 March 2011.

- Varfolomeev, E. E. *et al.* Targeted disruption of the mouse Caspase 8 gene ablates cell death induction by the TNF receptors, Fas/Apo1, and DR3 and is lethal prenatally. *Immunity* **9**, 267–276 (1998).
- Salmena, L. *et al.* Essential role for caspase 8 in T-cell homeostasis and T-cell-mediated immunity. *Genes Dev.* **17**, 883–895 (2003).
- Vercammen, D. *et al.* Inhibition of caspases increases the sensitivity of L929 cells to necrosis mediated by tumor necrosis factor. *J. Exp. Med.* **187**, 1477–1485 (1998).
- Holler, N. *et al.* Fas triggers an alternative, caspase-8-independent cell death pathway using the kinase RIP as effector molecule. *Nature Immunol.* **1**, 489–495 (2000).
- Wilson, N. S., Dixit, V. & Ashkenazi, A. Death receptor signal transducers: nodes of coordination in immune signaling networks. *Nature Immunol.* **10**, 348–355 (2009).
- Yeh, W. C. *et al.* FADD: essential for embryo development and signaling from some, but not all, inducers of apoptosis. *Science* **279**, 1954–1958 (1998).
- Yeh, W. C. *et al.* Requirement for Casper (c-FLIP) in regulation of death receptor-induced apoptosis and embryonic development. *Immunity* **12**, 633–642 (2000).
- Degterev, A. *et al.* Identification of RIP1 kinase as a specific cellular target of necrostatins. *Nature Chem. Biol.* **4**, 313–321 (2008).
- Zhang, D. W. *et al.* RIP3, an energy metabolism regulator that switches TNF-induced cell death from apoptosis to necrosis. *Science* **325**, 332–336 (2009).
- He, S. *et al.* Receptor interacting protein kinase-3 determines cellular necrotic response to TNF- α . *Cell* **137**, 1100–1111 (2009).
- Cho, Y. S. *et al.* Phosphorylation-driven assembly of the RIP1–RIP3 complex regulates programmed necrosis and virus-induced inflammation. *Cell* **137**, 1112–1123 (2009).
- Newton, K., Sun, X. & Dixit, V. M. Kinase RIP3 is dispensable for normal NF- κ Bs, signaling by the B-cell and T-cell receptors, tumor necrosis factor receptor 1, and Toll-like receptors 2 and 4. *Mol. Cell. Biol.* **24**, 1464–1469 (2004).
- Ogasawara, J. *et al.* Lethal effect of the anti-Fas antibody in mice. *Nature* **364**, 806–809 (1993).
- Siegel, R. M., Chan, F. K., Chun, H. J. & Lenardo, M. J. The multifaceted role of Fas signaling in immune cell homeostasis and autoimmunity. *Nature Immunol.* **1**, 469–474 (2000).
- Kang, T. B. *et al.* Mutation of a self-processing site in caspase-8 compromises its apoptotic but not its nonapoptotic functions in bacterial artificial chromosome-transgenic mice. *J. Immunol.* **181**, 2522–2532 (2008).
- Oberst, A. *et al.* Inducible dimerization and inducible cleavage reveal a requirement for both processes in caspase-8 activation. *J. Biol. Chem.* **285**, 16632–16642 (2010).
- Hughes, M. A. *et al.* Reconstitution of the death-inducing signaling complex reveals a substrate switch that determines CD95-mediated death or survival. *Mol. Cell* **35**, 265–279 (2009).
- Chang, D. W. *et al.* c-FLIP_L is a dual function regulator for caspase-8 activation and CD95-mediated apoptosis. *EMBO J.* **21**, 3704–3714 (2002).
- Boatright, K. M., Deis, C., Denault, J. B., Sutherland, D. P. & Salvesen, G. S. Activation of caspases-8 and -10 by FLIP_L. *Biochem. J.* **382**, 651–657 (2004).
- Chang, D. W. & Yang, X. Activation of procaspases by FK506 binding protein-mediated oligomerization. *Sci. STKE* **2003**, PL1 (2003).
- Zhou, Q. *et al.* Target protease specificity of the viral serpin CrmA. Analysis of five caspases. *J. Biol. Chem.* **272**, 7797–7800 (1997).
- Scaffidi, C. *et al.* Two CD95 (APO-1/Fas) signaling pathways. *EMBO J.* **17**, 1675–1687 (1998).
- Pop, C., Fitzgerald, P., Green, D. R. & Salvesen, G. S. Role of proteolysis in caspase-8 activation and stabilization. *Biochemistry* **46**, 4398–4407 (2007).
- Ch'en, I. L. *et al.* Antigen-mediated T cell expansion regulated by parallel pathways of death. *Proc. Natl Acad. Sci. USA* **105**, 17463–17468 (2008).
- Micheau, O., Lens, S., Gaide, O., Alevizopoulos, K. & Tschopp, J. NF- κ B signals induce the expression of c-FLIP. *Mol. Cell. Biol.* **21**, 5299–5305 (2001).
- Su, H. *et al.* Requirement for caspase-8 in NF- κ B activation by antigen receptor. *Science* **307**, 1465–1468 (2005).
- Scharner, D. *et al.* Caspase-8 is involved in neovascularization-promoting progenitor cell functions. *Arterioscler. Thromb. Vasc. Biol.* **29**, 571–578 (2009).
- Geserick, P. *et al.* Cellular IAPs inhibit a cryptic CD95-induced cell death by limiting RIP1 kinase recruitment. *J. Cell Biol.* **187**, 1037–1054 (2009).
- Smith, K. G., Strasser, A. & Vaux, D. L. CrmA expression in T lymphocytes of transgenic mice inhibits CD95 (Fas/APO-1)-transduced apoptosis, but does not cause lymphadenopathy or autoimmune disease. *EMBO J.* **15**, 5167–5176 (1996).
- Pop, C. *et al.* FLIP_L induces caspase-8 activity in the absence of interdomain caspase-8 cleavage and alters substrate specificity. *Biochem. J.* **433**, 447–457 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank W. J. Kaiser and E. S. Mocarski for their discussions and material support. We thank Ariad Pharmaceutical for providing the homo- and heterodimerization reagents. We thank the members of the St. Jude Immunology FACS core facility, as well as M. Johnson, L. Janke and the St. Jude Veterinary Pathology Core. We also thank the St. Jude Hartwell Center. This work was supported by NIH grant AI44828 to D.R.G. and CA69381 to G.S.S., as well as CIHR grant MOP 36537 to R.H. C.P.D. was supported by a fellowship grant from the Sass Foundation for Medical Research. This work was also supported by ALSAC.

Author Contributions A.O. and D.R.G. conceived the study and designed the experiments. C.P.D. and L.L.M. designed and conducted mouse breedings. C.P.D., R.W.

and A.O. carried out all experiments involving mice and tissues from mice. A.O. and R.W. carried out experiments involving cell lines and produced western blots. P.F. provided essential logistical and administrative support. C.P. and G.S.S. conceived, designed and carried out *in vitro* dimerization assays and CrmA inhibition studies. R.H. produced the caspase-8^{lox} animals that made the study possible.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.R.G. (douglas.green@stjude.org).

METHODS

siRNAs and DNA constructs. siRNAs against caspase-8 (5'-GUGAAUGGAAC CUGGUAUA-3' and 5'-GUCACGGACUUCAGACAAA-3', catalogue number J-043044-05 and J-043044-06, respectively) and FLIP (5'-GGAGCAAGAUU AAAUAUGA-3' and 5'-GAAUAGACUUGAACACAAA-3', catalogue number J-041091-05 and J-041091-08 respectively) as well as scramble siRNA control (5'-UGGUUUACAUGUCGACUAA-3', catalogue number D-001810-01) were ordered from Dharmacon, and introduced into cells using Lipofectamine RNAiMAX (Invitrogen) according to manufacturer's guidelines. Caspase-8 constructs were created by cloning full-length untagged murine caspase-8 upstream of a T2A ribosome-skipping sequence followed by green fluorescent protein (GFP), and introducing this construct into the pBabe-puro retroviral vector. The indicated D387A and C362A mutants were introduced using the QuikChange Mutagenesis kit from Agilent. For *in cellulo* experiments, CrmA was cloned into the pLNCX vector upstream of an IRES-GFP sequence, and cells transduced with this construct or vector control were sorted for GFP expression via FACS. Bcl-XL-expressing cell lines were produced by retroviral transduction of Bcl-XL-GFP fusion protein in pLZRS vector, followed by FACS sorting to achieve homogeneous expression.

Cell lines. L929, SVEC 4-10 and 3T3-SA cell lines stably expressing scrambled or RIPK3-specific short hairpin RNAs were produced using the TRCN0000022535 lentiviral construct (hairpin sequence: 5'-CCGGCCTCAGATTCCACATACTTT ACTCGAGTAAAG-TATGTGAATCTGAGGTTTTT-3') from OpenBiosystems or control shRNA (5'-CCTAAGGTAAAGTCGCCCTCGCTCGAGCGAGGGC GACTTAACCTTAGG-3') as previously described³¹. These cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% FCS, L-glutamine, and pen/strep. MEF were isolated from *Casp8^{-/-}Ripk3^{-/-}* embryos, transformed using E1A12S and H-RasG12V in pWZL-Hygro and pBabe-Puro retroviral vectors respectively, then selected in 0.5 $\mu\text{g ml}^{-1}$ puromycin and 40 $\mu\text{g ml}^{-1}$ hygromycin. These transformed cells were then transduced with the murine caspase-8 constructs described earlier, and sorted by FACS to achieve homogenous caspase-8 expression, then transduced with full-length untagged murine RIPK3 in the pLZRS retroviral vector and selected in 100 $\mu\text{g ml}^{-1}$ zeocin. MEF were maintained in DMEM as described above, but also supplemented with 55 μM β -mercaptoethanol, 1 mM sodium pyruvate and non-essential amino acids (Gibco). Caspase-8^{fllox/fllox}RosaCreER MEF were transformed as described above, and Cre recombinase was activated by culturing cells in 100 nM 4-hydroxytamoxifen for 48 h, followed by culture for 1 week to establish a stable population. Caspase-8 deletion was confirmed by PCR and western blot (not shown).

Immunoprecipitation of FADD. 3T3-SA cells stably overexpressing Bcl-XL were transfected with siRNAs specific to FLIP, caspase-8, or with a scrambled control siRNA as described. 48 h post-transfection, cells were treated with 20 ng ml⁻¹ recombinant murine TNF- α for 90 min. Immunoprecipitation of DISC-associated complexes was carried out using buffer and lysis conditions previously described³². FADD was immunoprecipitated using the M19 polyclonal anti-FADD antibody conjugated to Protein A/G-PLUS Agarose beads, also from Santa Cruz. Immune complexes were eluted by boiling in reducing western blot loading buffer and resolved by western blot using the antibodies described.

Immune cell staining, cell death and activation assays. For immune cell staining, spleen, thymus and lymph node were harvested from animals and single cell suspensions were generated. For immune cells staining from the blood, blood was harvested weekly from the retroorbital sinus from animals anesthetized with 2–2.5% isoflurane in 11 oxygen. Red blood cells were lysed in hypotonic buffer and samples were stained with the appropriate antibodies as described below. Data was acquired using a FACsCalibur or LSR II using CellQuest Pro and FACsDiva software, respectively. Data analysis was performed using FlowJo (Tree Star). For activation assays, splenic T cells were isolated from whole spleen using magnetic separation (Pan T cell isolation kit II, Miltenyi Biotec (130-095-130)). Cells were stained with CFSE and plated at 1×10^5 cells per well in 96-well plates with 10 $\mu\text{g ml}^{-1}$ plate-bound anti-CD3 and 10 $\mu\text{g ml}^{-1}$ soluble anti-CD28. For thymocyte death assays, single cell suspensions of thymocytes were plated at 1.5×10^5

cells per well in 96-well plates and treated with the various apoptosis inducers with or without 20 μM qVD. Cells were harvested at 8 h (or 22 h for Jo2), stained with AnnexinV and 7-AAD, and assayed for viability using flow cytometry. Splenocytes were cultured in RPMI 1640 (Invitrogen) supplemented as described for MEF above. Thymocytes were cultured in this media as well, but charcoal-stripped FBS was used.

Compounds, antibodies and cytokines. Antibodies used for western blot were: Anti-RIPK1 from BD (610458), anti-RIPK3 from Imgenex (IMG-5523-2), anti-caspase-8 (1G12) and anti-FLIP (Dave-2) both from Alexis, and anti-Actin (C4) from MP, anti-caspase-3 from Cell Signaling (9662), and anti-Bcl-XL(S18) and anti-FADD(M19) both from Santa Cruz. Murine TNF- α came from Peprotech. Necrostatin-1 came from Enzo, and zVAD-fmk came from SM Biochemicals. Purified anti-murine CD95 (clone Jo2), antiCD3 (clone 145-2C11), and anti-CD28 (clone 37.51), and fluorescent-conjugated anti-B220-PE (clone RA3-6B2), anti-CD4-PerCp-Cy5.5 (clone RM4-5), anti-V β 6-FITC and anti-V β 8-PE came from BD Biosciences. Anti-CD3-FITC (clone 145-2C11) and anti-CD8-APC (clone 53-6.7), as well as 7-AAD came from eBioscience. Annexin-V-APC came from Invitrogen. Dexamethasone was from APP Pharmaceuticals, staurosporine, etoposide and ionomycin were from Sigma. Ultraviolet irradiation was accomplished using a UV Stratalinker 2400 from Stratagene, whereas γ -irradiation used a JL Shepard Mark 1 Cesium 137 irradiator.

Protein purification and *in vitro* assays. Production of CrmA³³ and FKBP-caspase-8 (ref. 34) were accomplished as previously described. FRB-FLIP_L mutants were expressed as C-terminal His-tagged proteins in pET29b. Upon induction with 1 mM IPTG, cells were grown at 25 °C for 16 h and purified by Ni-affinity chromatography. Homodimerization and kosmotrope assays have been described³⁴. Heterodimerization experiments were performed similarly, except that FRB-FLIP_L and the heterodimerization drug AP21967 (provided, along with the homodimerization drug AP20187, by Ariad Pharmaceuticals, <http://www.ariad.com/regulationkits>) were added in 4–5-fold molar excess to FKBP-caspase-8. For CrmA inhibition studies, homo or heterodimeric activated caspase-8 (10–50 nM) diluted in caspase buffer (10 mM Pipes pH 7.2, 0.1 M NaCl, 1 mM EDTA, 10% sucrose, 0.05% CHAPS, 5 mM DTT) was added to a mixture of serially diluted CrmA (30–2,000 nM). Caspase-8 activity was monitored using Ac-IETD-AFC; indicated activities represent linear rates of IETDase activity between 850 and 950 s after addition of CrmA, as a percentage of activity without CrmA addition.

Mice and treatments. Mice with a deleted allele of caspase-8 were generated by germ line deletion of a caspase-8^{fllox} allele described previously³⁵. RIPK3-deficient animals were obtained from V. Dixit³⁶. Genotypes were confirmed by tail snip PCR as described previously. For Jo2 injections, animals were injected via tail vein with 15 μg purified Jo2 in lipopolysaccharide-free PBS per animal. Liver enzymes were assayed using a Trilogy Multi-Purpose Analyzer System from Drew Scientific, and liver sections were created and stained with haematoxylin and eosin, in the St. Jude Veterinary Pathology Core facility. For SEB injections, 50 μg SEB (Toxin Technology Inc.) per animal was injected via tail vein and T-cell populations were monitored by retro-orbital bleed and FACS as detailed previously. The St. Jude Institutional Animal Care and Use Committee approved all procedures in accordance with the Guide for the Care and Use of Animals.

- Upton, J. W., Kaiser, W. J. & Mocarski, E. S. Virus inhibition of RIP3-dependent necrosis. *Cell Host Microbe* **7**, 302–313 (2010).
- Geserick, P. *et al.* Cellular IAPs inhibit a cryptic CD95-induced cell death by limiting RIP1 kinase recruitment. *J. Cell Biol.* **187**, 1037–1054 (2009).
- Quan, L. T., Caputo, A., Bleackley, R. C., Pickup, D. J. & Salvesen, G. S. Granzyme B is inhibited by the cowpox virus serpin cytokine response modifier A. *J. Biol. Chem.* **270**, 10377–10379 (1995).
- Oberst, A. *et al.* Inducible dimerization and inducible cleavage reveal a requirement for both processes in caspase-8 activation. *J. Biol. Chem.* **285**, 16632–16642 (2010).
- Salmena, L. *et al.* Essential role for caspase 8 in T-cell homeostasis and T-cell-mediated immunity. *Genes Dev.* **17**, 883–895 (2003).
- Newton, K., Sun, X. & Dixit, V. M. Kinase RIP3 is dispensable for normal NF- κ Bs, signaling by the B-cell and T-cell receptors, tumor necrosis factor receptor 1, and Toll-like receptors 2 and 4. *Mol. Cell. Biol.* **24**, 1464–1469 (2004).

RIP3 mediates the embryonic lethality of caspase-8-deficient mice

William J. Kaiser¹, Jason W. Upton¹, Alyssa B. Long², Devon Livingston-Rosanoff¹, Lisa P. Daley-Bauer¹, Razqallah Hakem³, Tamara Caspary² & Edward S. Mocarski¹

Apoptosis and necroptosis are complementary pathways controlled by common signalling adaptors, kinases and proteases; among these, caspase-8 (Casp8) is critical for death receptor-induced apoptosis. This caspase has also been implicated in non-apoptotic pathways that regulate Fas-associated via death domain (FADD)-dependent signalling and other less defined biological processes as diverse as innate immune signalling and myeloid or lymphoid differentiation patterns¹. Casp8 suppresses RIP3–RIP1 (also known as RIPK3–RIPK1) kinase complex-dependent^{2–4} necroptosis⁵ that follows death receptor activation as well as a RIP3-dependent, RIP1-independent necrotic pathway that has emerged as a host defence mechanism against murine cytomegalovirus⁶. Disruption of *Casp8* expression leads to embryonic lethality in mice between embryonic days 10.5 and 11.5 (ref. 7). Thus, Casp8 may naturally hold alternative RIP3-dependent death pathways in check in addition to promoting apoptosis. We find that RIP3 is responsible for the mid-gestational death of Casp8-deficient embryos. Remarkably, *Casp8*^{−/−}*Rip3*^{−/−} double mutant mice are viable and mature into fertile adults with a full immune complement of myeloid and lymphoid cell types. These mice seem immunocompetent but develop lymphadenopathy by four months of age marked by accumulation of abnormal T cells in the periphery, a phenotype reminiscent of mice with Fas-deficiency (*lpr/lpr*; also known as *Fas*). Thus, Casp8 contributes to homeostatic control in the adult immune system; however, RIP3 and Casp8 are together completely dispensable for mammalian development.

To determine whether Casp8 can hold RIP3 kinase-dependent death^{2–5,8} in check, we used murine L929 cells, a system that requires continued Casp8 expression for cell survival⁹. Inhibition of Casp8 with either small interfering RNA (siRNA; Supplementary Fig. 1a) or zVAD-fmk (Supplementary Fig. 1c) induced death, as expected from previous studies⁹. When treated with RIP3-specific short hairpin RNA (shRNA), however, L929 cells were protected from death (Supplementary Fig. 1a), consistent with this being a RIP3-dependent necrotic death pathway. The murine cytomegalovirus (MCMV) M45 gene-encoded viral inhibitor of RIP activation (vIRA) blocks RIP3-dependent necrotic death⁶. In keeping with the importance of a RIP3–RIP1 complex in necroptosis^{2–4}, vIRA blocked death, whereas a tetra-alanine RIP homotypic interaction motif (RHIM) substitution mutant, M45mutRHIM⁶ failed to suppress death induced by either Casp8 siRNA or zVAD (Supplementary Figs 1b, d). Necrostatin-1 was used to demonstrate that RIP1 kinase activity was necessary for necroptosis (Supplementary Fig. 1c). The specific viral inhibitor of Casp8 activation (vICA) encoded by the MCMV M36 gene¹⁰ also induced this death pathway (Supplementary Fig. 1e, f). These data demonstrate that L929 cells succumb to RIP3-dependent necrotic death when Casp8 is inactive or eliminated.

Disruption of *Casp8* expression leads to embryonic lethality in mice between embryonic day 10.5 (E10.5) and E11.5, coincident with embryonic vascular, cardiac and haematopoietic defects^{7,11–13}; however,

the molecular mechanisms behind these defects remain poorly defined^{7,11–13}. To evaluate the potential contribution of RIP3 to embryonic lethality, we examined wild-type and *Casp8*^{−/−} embryos as well as extra-embryonic tissues for *Rip3* expression. On the basis of *in situ* hybridization, *Rip3* transcript levels increased and tissue distribution broadened as embryonic development proceeded from E9.5 to E12.5 in both genotypes, indicating that *Rip3* transcript is not regulated by *Casp8* (Fig. 1a and data not shown). At E9.5, *Rip3* was prominent in the apical ectodermal ridge (AER) of the hind-limb bud and the tail bud, and expanded to include the fore-limb bud AER, midline of the spinal cord, branchial arches and intersomitic regions by E10.5 through E12.5 (Fig. 1a and data not shown), a signal that was absent from RIP3-deficient E10.5 embryos (Fig. 1b). Additionally, RIP3 was readily detected by immunoblot in E10.5 embryo and yolk sac cell lysates (data not shown). Mice homozygous for disruption of exons 3 and 4 in *Casp8* (ref. 14) died by E11.5 and presented with heart defects, hyperaemia in the abdominal region, and undulation of the neural tube (Fig. 1c and Supplementary Fig. 2a), consistent with earlier studies that used alternative targeting strategies to disrupt *Casp8* (refs 7, 11, 13).

The yolk sac is the initial site of haematopoiesis before the transfer of function to intra-embryonic sites. The most pronounced impact of *Casp8* disruption, either in embryonic tissues or in the yolk sac, is the disruption of endothelial cell organization leading to circulatory failure in the yolk sac, the probable culprit behind embryonic lethality^{11,13}. When we used PECAM-specific antibody to localize endothelial cells in *Casp8*-deficient embryos (E10.5), we observed the expected contrast between an organized yolk sac vascular pattern in control mice and a disrupted pattern in *Casp8*-deficient mice (Fig. 1d). These observations affirmed the disruption of vascular development in *Casp8* null embryos. RIP3 was also detected in the yolk sac endothelium of both *Casp8*^{+/−}*Rip3*^{+/−} and *Casp8*^{−/−}*Rip3*^{+/−} embryos at this time (Supplementary Fig. 2b); thus, RIP3 was present in the cell populations most widely implicated in embryonic death of *Casp8*-deficient mice independently of *Casp8* expression. *Casp8* deficiency is also known to compromise primitive haematopoietic progenitor cell (HPC) development⁷, a process dependent on CD41⁺ cells populating yolk sac blood islands and fetal sites¹⁵. RIP3 was detected in CD41⁺ cells in the yolk sac blood islands of embryos (Fig. 1e and Supplementary Fig. 2c). The striking CD41⁺ cell fragmentation in the yolk sac blood islands of *Casp8*^{−/−}*Rip3*^{+/−} embryos at E10.5 further implicates this kinase in processes leading to embryonic death.

To establish the role of the RIP3 kinase in embryonic lethality of *Casp8*-deficient mice, we generated double knock out (DKO) *Casp8*^{−/−}*Rip3*^{−/−} embryos by a *Casp8*^{+/−}*Rip3*^{+/−} intercross. In contrast to *Casp8*^{−/−}*Rip3*^{+/−} embryos, which were developmentally arrested at ~E11.0, DKO embryos were indistinguishable from *Casp8*^{+/−}*Rip3*^{−/−} or *Casp8*^{+/−}*Rip3*^{+/−} embryos at E12.5 and later times, had a functioning heart and organized yolk sac endothelial architecture and, remarkably, exhibited HPC colony formation at levels comparable to wild-type mice (Fig. 1f, g and Supplementary

¹Department of Microbiology and Immunology, Emory Vaccine Center, Emory University School of Medicine, Atlanta, Georgia 30322, USA. ²Department of Human Genetics, Emory University, Atlanta, Georgia 30322, USA. ³Department of Medical Biophysics, University of Toronto and Ontario Cancer Institute, University Health Network, Toronto, Ontario M5G 2M9, Canada.

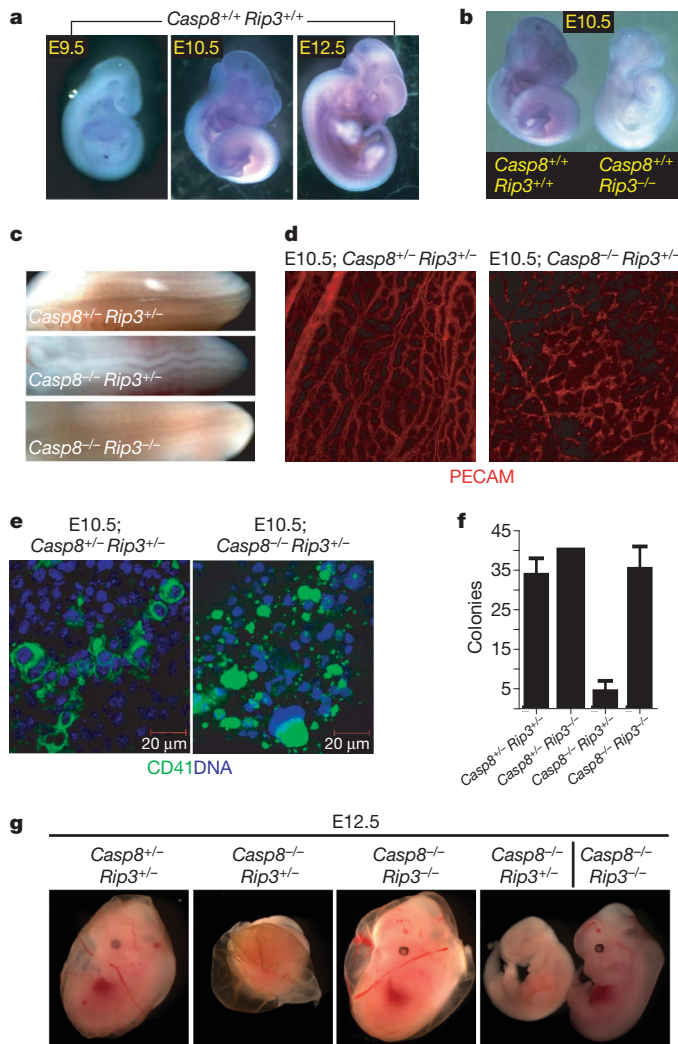


Figure 1 | Embryonic expression of *Rip3*. **a**, Whole-mount *Rip3* *in situ* hybridization of *Casp8*^{+/+}*Rip3*^{+/+} E9.5 (left panel), E10.5 (middle panel), and E12.5 (right panel) embryos. **b**, Whole-mount *Rip3* *in situ* hybridization of *Casp8*^{+/+}*Rip3*^{+/+} and *Casp8*^{+/+}*Rip3*^{-/-} E10.5 embryos demonstrating specificity of the probe. **c**, View of the neural tube of E11.5 embryos with the indicated genotype. **d**, PECAM-1 (CD31) staining of a whole-mount E10.5 yolk sac from a representative *Casp8*^{+/+}*Rip3*^{-/-} (left panel) and *Casp8*^{-/-}*Rip3*^{+/+} (right panel) embryo (×100). **e**, CD41 (green) and nuclear DNA (blue) staining of a yolk sac from a E10.5 *Casp8*^{+/+}*Rip3*^{+/+} (left panel) and a *Casp8*^{-/-}*Rip3*^{+/+} (right panel) embryo. **f**, Average number of colony-forming cells (CFC) following culture of disrupted E10.5 yolk sacs of the indicated genotype. Error bars, s.d. (*n* = 3). **g**, Photographs of E12.5 embryos and yolk sacs with the indicated genotype. The right panel shows side by side embryos with yolk sacs removed.

Fig. 2d and data not shown). The apparent normalization of the *Casp8*-deficient phenotype by removal of *Rip3* indicated that this kinase was responsible for the ~E11.0 embryonic block, so we permitted the intercross pregnancies to complete gestation. PCR analysis performed on tissues from weanling mice confirmed deletion of *Casp8* exons 3–4 as well as deletion of *Rip3* in both alleles (Supplemental Fig. 3a). We detected *Rip1*, *Casp8* and *Rip3* in spleen and thymus from wild-type mice examined by immunoblot (Fig. 2a). *Casp8*^{+/+}*Rip3*^{-/-}, *Casp8*^{+/+}*Rip3*^{-/-} and DKO mice all lacked *Rip3* but retained unaltered *Rip1* levels in tissues. *Casp8* was absent from DKO mice but present in other intercross progeny with predicted Mendelian frequencies (Fig. 2b). As expected, *Casp8*-deficient, *Rip3*-expressing progeny were not seen. Furthermore, when bred, adult DKO mice gave birth to viable mice that survived through adulthood and appeared similar

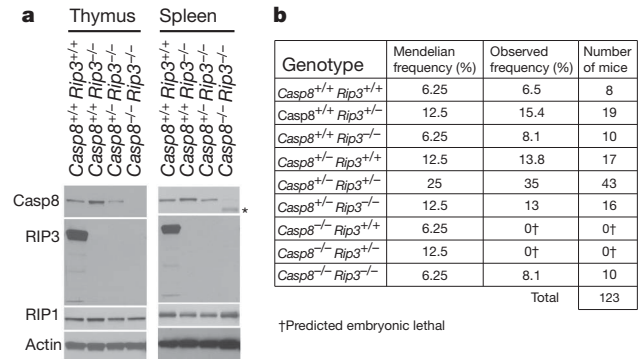


Figure 2 | *Casp8*^{-/-}*Rip3*^{-/-} mice are viable. **a**, Immunoblot of *Casp8*, *Rip3*, *Rip1* and β -actin from thymus (left panel) and spleen (right panel). The asterisk denotes elevated heavy IgG heavy chain reactive with secondary antibody in DKO sample. **b**, Epistasis analysis of mice born following *Casp8*^{+/+}*Rip3*^{+/+} intercross with predicted and observed frequencies.

to *Casp8*^{+/+}*Rip3*^{-/-} mice bred in parallel (Supplementary Fig. 3b). Earlier studies showed that conditional deletion of *Casp8* in epidermis promoted perinatal lethality characterized by chronic inflammation^{16,17}. DKO mice did not exhibit any evidence of skin inflammation when followed for more than 6 months, indicating that *Rip3* must have had a proinflammatory role in the absence of epithelial *Casp8* in prior studies. These results further indicate that embryonic lethality as well as a range of vascular degeneration and neural tube defects, together with hematopoietic abnormalities and chronic inflammation seen in mice lacking *Casp8* are all *Rip3*-dependent.

To verify the functional elimination of *Casp8* in cells derived from DKO animals, we evaluated susceptibility to inducers of death receptor (DR)-mediated apoptosis. Although *Casp8* has an essential role in macrophage differentiation¹¹, DKO mice produced CD11b⁺F4/80⁺ bone marrow-derived mononuclear (BMDM) cells just as readily as controls, and these cells lacked detectable *Rip3* or *Casp8* (Fig. 3a and b). Thus, the requirement for *Casp8* during macrophage differentiation is suppressed by *Rip3*-deficiency. Macrophages prepared from wild-type, *Casp8*^{+/+}*Rip3*^{-/-} and *Casp8*^{+/+}*Rip3*^{-/-} mice died when exposed to reagents promoting Fas activation, whereas DKO cells were completely resistant, consistent with the established role of this caspase in DR-induced apoptosis. In the presence of the caspase inhibitor zVAD-fmk, Fas activation promoted *Rip3*-dependent necroptosis in *Rip3*-containing BMDM cells (Fig. 3c). Furthermore, cells expressing *Rip3* were susceptible to necroptosis, whereas *Rip3* knockout and DKO mice remained resistant to this death pathway. Thus, DKO cells were insensitive to either extrinsic DR-induced apoptosis or *Rip3*-dependent necroptosis, consistent with their genotype. Previously, conditional knockout of *Casp8* in the liver revealed its essential role in TNF or Fas-induced fatal hepatitis^{11,18}. DKO mice were resistant to anti-Fas-antibody treatment, survived for over 48 h (Fig. 3d) and showed normal liver architecture (Fig. 3e), although they exhibited slightly elevated levels of the liver-associated transaminases ALT (alanine aminotransferase) and AST (aspartate aminotransferase) (data not shown). In contrast, *Casp8*^{+/+}*Rip3*^{-/-} littermate control mice developed hepatitis (Fig. 3d and 3e). DKO mice were also resistant to administration of LPS in combination with the liver-specific transcriptional inhibitor D-(+)-galactosamine (GalN), which induces a TNF-dependent fatal liver hepatitis in both wild-type¹⁸ and *Casp8*^{+/+}*Rip3*^{-/-} mice (Supplementary Fig. 4). This resistance was comparable to the negative control TRIF-deficient mice¹⁹ (also known as TICAM1; Supplementary Fig. 4).

Although *Rip3* is dispensable for myeloid and lymphoid development²⁰, *Casp8* is essential for generation of both myeloid and lymphoid lineages^{11,14,21,22}. To determine whether this essential role of *Casp8* was due to dysregulation of *Rip3*, we evaluated the characteristics of leukocytes from the thymus, bone marrow, spleen and lymph node

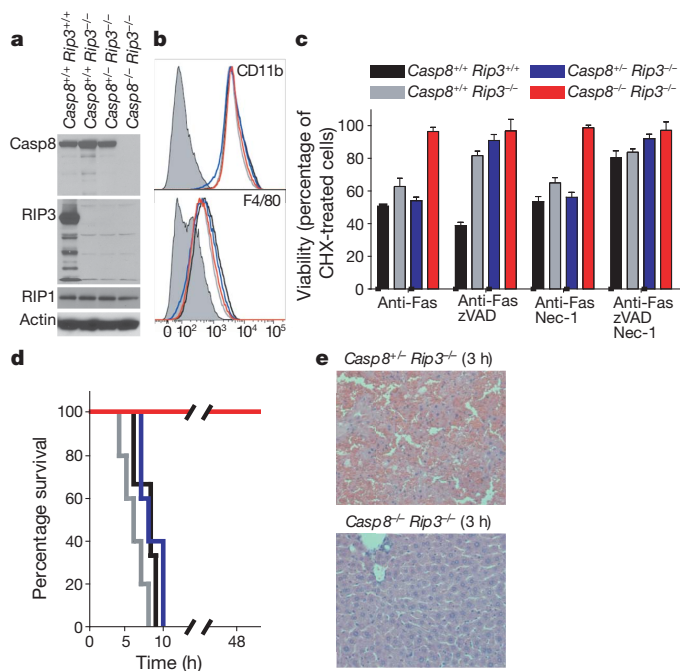


Figure 3 | Sensitivity to DR-induced apoptosis, necroptosis and disease.

a, Immunoblot of Casp8, RIP3, RIP1 and β -actin in BMDM derived from mice with the indicated genotype. **b**, Relative cell surface expression levels of CD11b (top panel) or F4/80 (bottom panel) shown by red line (DKO), black line ($Casp8^{+/-}Rip3^{+/+}$), grey line ($Casp8^{+/-}Rip3^{-/-}$), or blue line ($Casp8^{+/-}Rip3^{-/-}$) on BMDM stained cells. Isotype control is shown by shaded histogram. **c**, Viability of BMDM cultured in CHX ($5 \mu\text{g ml}^{-1}$) and treated with anti-Fas antibody for 18 h in the presence or absence of the caspase inhibitor zVAD-fmk ($25 \mu\text{M}$) and/or Nec-1 ($30 \mu\text{M}$). Cell viability was determined by measuring intracellular ATP levels with a Cell Titer-Glo Luminescent Cell Viability Assay kit. Error bars, s.d. ($n = 4$). **d**, Kaplan-Meier survival plot of 12-week-old $Casp8^{+/-}Rip3^{+/+}$ ($n = 3$), $Casp8^{+/-}Rip3^{-/-}$ ($n = 5$), $Casp8^{+/-}Rip3^{-/-}$ ($n = 5$), and DKO ($n = 5$) mice injected intraperitoneally with $12.5 \mu\text{g}$ of anti-Fas Jo-2 antibody; n , numbers of mice analysed. Legend genotypes and line colour is the same as in **c**. **e**, Histology of liver sections from $Casp8^{+/-}Rip3^{-/-}$ (left panel) and $Casp8^{-/-}Rip3^{-/-}$ mice (right panel) 3 h after injection with Jo-2 antibody.

(LN) of DKO mice, littermate $Casp8^{+/-}Rip3^{-/-}$ and wild-type mice (Fig. 4a) by fluorescence-activated cell sorting (FACS) analysis, detecting the presence of myeloid and lymphoid populations in all three genotypes based on forward (size) and side (granularity) light scatter properties (data not shown) in combination with surface markers. We observed inflammatory monocytes ($CD11b^{+}Ly6C^{hi}$) and polymorphonuclear leukocytes ($CD11b^{+}Ly6C^{int}$) in the bone marrow and spleen (Fig. 4a) of wild-type and DKO mice. On the basis of CD11c and F4/80 expression, dendritic cells and macrophages, respectively, populated the same tissues of DKO and $Casp8^{+/-}Rip3^{-/-}$ littermate controls (data not shown). Thus, myeloid cell populations continued to be generated in the absence of RIP3 and Casp8, consistent with the successful derivation of BMDM cells from DKO mice (see Fig. 3). NK ($CD49b^{+}CD3^{-}$), NKT ($CD49b^{+}CD3^{+}$) and B ($CD19^{+}$) lymphocytes, including predominant IgD^{+} as well as less prevalent IgD^{-} B cells (data not shown) were present in all tissues examined (Fig. 4a).

Defects in the Fas (CD95)-death receptor pathway promote the accumulation and expansion of lymphocytes and the development of autoimmune lymphoproliferative syndromes (ALPS)²³. Caspase-8 is downstream of Fas, and similarly DKO mice showed pronounced splenomegaly and lymphadenopathy over the first few months of age (Supplementary Fig. 5a–c). Adult DKO mouse spleens ranged from three to seven times the size of $Casp8^{+/-}Rip3^{-/-}$ littermate controls (Fig. 4b and Supplementary Fig. 5b, c) and contained more lymphoid cells in splenic white pulp (data not shown). Histological examination

revealed lymphocytic infiltrates in the salivary glands, pancreas and lamina propria of both stomach and small intestine (data not shown). Consistent with size, DKO mice had significantly greater numbers of leukocytes in secondary lymphoid tissues such as spleen (Supplementary Fig. 5d) that seemed to result from abnormally high levels of $CD3^{+}$ T cells and to a lesser degree $CD19^{+}$ B cells (Supplementary Fig. 5e). Whereas these characteristics are consistent with the known role of Casp8 in DR-associated haematopoietic homeostasis²³, they contrast with the characteristics of mice with Casp8-deficient T cells, where there are fewer T cells than B cells in secondary lymphoid tissues^{14,22}. Thus, there was a marked accumulation of T cells that contributed to lymphadenopathy and splenomegaly as DKO mice aged.

Casp8-deficient T cells show defects in the response to antigens and mitogens^{14,24}. To compare T-cell activation in DKO and $Casp8^{+/-}Rip3^{-/-}$ littermate controls, we evaluated the sensitivity of bulk splenocytes from mock and MCMV-infected mice, to anti-CD3 and anti-CD28 treatment. In contrast to earlier observations on Casp8-deficient T cells^{14,24}, this treatment induced a response in mock-infected DKO cells and this response was enhanced at 7 days post MCMV infection. In fact, the CD8 T cells from DKO mice responded with an increased frequency in bifunctional $\text{INF}\gamma^{+}\text{TNF}^{+}$ cells (Fig. 4c). Furthermore, DKO mice survived a dose of virus that is lethal to immunocompromised *scid/scid* (also known as *Prkdc*) mice and maintained control over viral replication levels for 30 days, similar to $Casp8^{+/-}Rip3^{-/-}$ littermate controls (data not shown). Although more characterization is required to understand fully the quality of the immune response in DKO mice, T cell receptor-dependent activation of naive and enhanced activation of antigen-exposed T cells is clearly retained despite the combined disruption of *Casp8* and *Rip3*.

Fas-deficient (*lpr/lpr*) mice are marked by the accumulation in the periphery of the $CD3^{+}$ T cells that are $B220^{+}$ but lack CD4 and CD8; these cells have been ascribed to a failure of apoptosis²⁵. DKO mice had normal CD4 and CD8 T-cell populations in the thymus (Supplementary Fig. 5f), but contained this signature $B220^{+}$ T cell population in secondary lymphoid organs (Fig. 4d and Supplementary Fig. 5g). Interestingly, conditional deletion of Casp8 in T cells, although resulting in lymphadenopathy, did not result in the emergence of this signature phenotype²², indicating that RIP3-dependent necroptosis eliminates T cells where Casp8 is absent or non-functional. Thus, Fas DR-induced death pathways are essential for immune homeostasis and in the absence of both Casp8 and RIP3, excess unusual T cells probably accumulate owing to a failure of both apoptosis and necroptosis.

The data presented here point to RIP3 as the vital target of Casp8 during mammalian development, with Casp8 functioning during embryogenesis to restrict RIP3 rather than to mediate apoptosis. We speculate that the RIP3-dependent pathway controlled by Casp8 in the vasculature and haematopoietic cells during embryogenesis, as well as in other tissues where Casp8 deficiency leads to cell death, is related to known necrotic pathways controlled by this kinase. Eliminating this RIP3-dependent embryonic death in combination with Casp8-deficiency results in the accumulation of abnormal T cells as adult animals age, but does not lead to defects at other developmental stages. Although Casp8 may have other non-apoptotic roles such as influencing cell motility and proliferation, one of its key roles is its essential non-apoptotic function to control RIP3. FADD promotes Casp8 activation within the death-inducing signalling complex (DISC), a process subject to regulation by cFLIP. Deletion of FADD, cFLIP or Casp8 in all tissues causes a common pattern of vascular and haematopoietic defects associated with embryonic lethality in mice around E10.5 (ref. 1), suggesting the coordinated control of RIP3 by these three players during development. The signals controlling activation of Casp8 and/or RIP3 during embryogenesis remain to be identified. Nevertheless, therapeutic strategies targeting DISC components aimed at apoptosis will need to take into account the likely triggering of RIP3-dependent pathways. Upon combined elimination of Casp8 and RIP3,

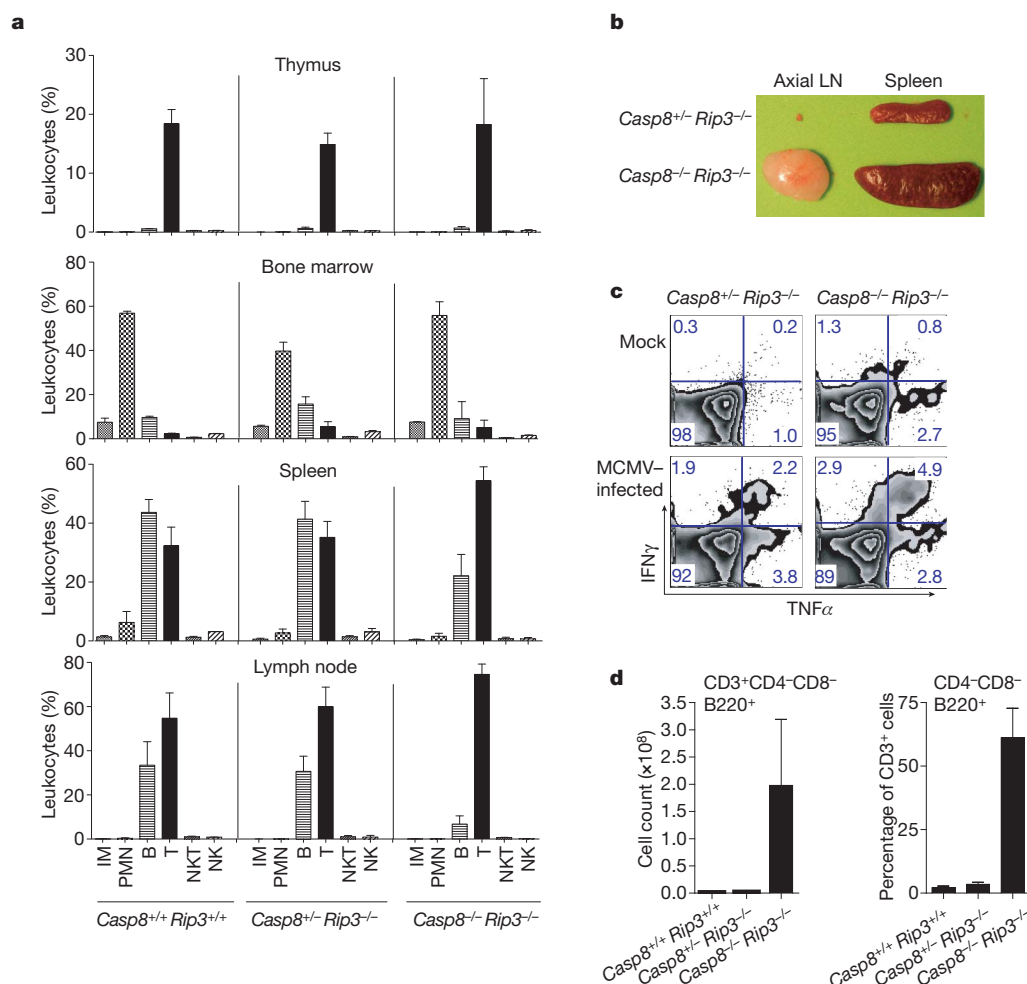


Figure 4 | Immune compartment of 16-week-old DKO mice. **a**, Live cells from thymus (top panels), bone marrow (second set of panels), spleen (third set of panels) and LN (bottom panels), gated based on forward and side scatter properties, and stained for surface expression of CD19, CD3, CD49b, Ly6C and CD11b to define non-overlapping leukocyte (CD45⁺) populations. The average and s.d. for three wild-type (left panels), four *Casp8*^{+/+}*Rip3*^{-/-} (middle panels) littermate control and three DKO (right panels) mice showing levels of inflammatory monocytes (IM), polymorphonuclear leukocytes

(PMN), B cells, T cells, NK cells and NKT cells. **b**, Photograph of the axial LN and spleen from representative mice of the indicated genotype. **c**, Flow cytometric analysis of IFN γ and TNF α on T cells from naive or MCMV-infected (7 days) mice following stimulation with CD3/CD28 antibodies. CD8⁺ T cells from one representative animal per experimental group are shown. **d**, B220 expression was assessed on CD4⁻CD8⁻ splenic T cells. Error bars, s.d. for three wild-type, four *Casp8*^{+/+}*Rip3*^{-/-} littermate control and three DKO mice.

defects in DR-dependent lymphocyte differentiation and homeostasis emerge. The discovery that Casp8 suppresses RIP3 pathways and is essential for maintenance of the vasculature, haematopoiesis, suppression of the innate immune system and T-cell function points to the unexpected importance of these pathways in humans and other mammals. The observation here that Casp8 is a gatekeeper, suppressing RIP3 during development in addition to promoting apoptosis mediated via death receptors, promises to have broad implications in approaches to cancer therapy and regenerative medicine as well as during elaboration of the innate and adaptive phases of the immune response.

METHODS SUMMARY

RIP3^{-/-} (ref. 20) and TRIF mutant (strain C57BL/6J-Ticam1^{Lps2})¹⁹ mice have been described previously. *Casp8*^{+/+} mice were generated by crossing *Casp8*^{fl3-4/wt} (ref. 14) with Rosa-CreER mice²⁶ *Rip3*^{-/-} and *Casp8*^{+/+} mice were subsequently intercrossed. Embryos and yolk sacs were harvested from timed pregnancies at the indicated day post-coitus. Genotypes were determined by PCR from tail-snips or fetal tissue as described. Fatal hepatitis was induced by intraperitoneal injection with 12.5 μ g of anti-mouse Fas antibody (clone Jo2, BD Biosciences). Tissue processing and staining was performed by Emory University Division of Animal Resources (EU-DAR). Mice were bred and maintained by EU-DAR, where all procedures were approved by the Emory University Institutional

Animal Care and Use Committee. Immunoblotting, preparation of protein extracts, and immunoprecipitations were as described previously²⁷. Whole-mount *in situ* hybridization was performed as described²⁸, with digoxigenin-labelled antisense RNA probes transcribed from linearized RIP3-encoding plasmid (accession number BC029210, ATCC) according to the manufacturer's directions (Roche). Yolk sacs were fixed and stained for immunofluorescence microscopy by standard methods, and images were acquired on a Carl Zeiss LSM 510 META confocal fluorescence microscope. Bone-marrow-derived macrophages were generated and viability determined as described previously⁶. Cells for flow cytometry were harvested, processed and stained with indicated antibodies by standard methods. Data were acquired using an LSRII flow cytometer (BD Biosciences) and analysed with FlowJo software.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 November 2010; accepted 24 January 2011.

Published online 2 March 2011.

1. Strasser, A., Jost, P. J. & Nagata, S. The many roles of FAS receptor signaling in the immune system. *Immunity* **30**, 180–192 (2009).
2. Cho, Y. S. *et al.* Phosphorylation-driven assembly of the RIP1–RIP3 complex regulates programmed necrosis and virus-induced inflammation. *Cell* **137**, 1112–1123 (2009).
3. He, S. *et al.* Receptor interacting protein kinase-3 determines cellular necrotic response to TNF- α . *Cell* **137**, 1100–1111 (2009).

4. Zhang, D. W. *et al.* RIP3, an energy metabolism regulator that switches TNF-induced cell death from apoptosis to necrosis. *Science* **325**, 332–336 (2009).
5. Holler, N. *et al.* Fas triggers an alternative, caspase-8-independent cell death pathway using the kinase RIP as effector molecule. *Nature Immunol.* **1**, 489–495 (2000).
6. Upton, J. W., Kaiser, W. J. & Mocarski, E. S. Virus inhibition of RIP3-dependent necrosis. *Cell Host Microbe* **7**, 302–313 (2010).
7. Varfolomeev, E. E. *et al.* Targeted disruption of the mouse Caspase 8 gene ablates cell death induction by the TNF receptors, Fas/Apo1, and DR3 and is lethal prenatally. *Immunity* **9**, 267–276 (1998).
8. Chan, F. K. *et al.* A role for tumor necrosis factor receptor-2 and receptor-interacting protein in programmed necrosis and antiviral responses. *J. Biol. Chem.* **278**, 51613–51621 (2003).
9. Yu, L. *et al.* Regulation of an ATG7-*beclin 1* program of autophagic cell death by caspase-8. *Science* **304**, 1500–1502 (2004).
10. McCormick, A. L., Skaletskaya, A., Barry, P. A., Mocarski, E. S. & Goldmacher, V. S. Differential function and expression of the viral inhibitor of caspase 8-induced apoptosis (vICA) and the viral mitochondria-localized inhibitor of apoptosis (vMIA) cell death suppressors conserved in primate and rodent cytomegaloviruses. *Virology* **316**, 221–233 (2003).
11. Kang, T. B. *et al.* Caspase-8 serves both apoptotic and nonapoptotic roles. *J. Immunol.* **173**, 2976–2984 (2004).
12. Kang, T. B. *et al.* Mutation of a self-processing site in caspase-8 compromises its apoptotic but not its nonapoptotic functions in bacterial artificial chromosome-transgenic mice. *J. Immunol.* **181**, 2522–2532 (2008).
13. Sakamaki, K. *et al.* *Ex vivo* whole-embryo culture of caspase-8-deficient embryos normalize their aberrant phenotypes in the developing neural tube and heart. *Cell Death Differ.* **9**, 1196–1206 (2002).
14. Salmena, L. *et al.* Essential role for caspase 8 in T-cell homeostasis and T-cell-mediated immunity. *Genes Dev.* **17**, 883–895 (2003).
15. Mitjavila-Garcia, M. T. *et al.* Expression of CD41 on hematopoietic progenitors derived from embryonic hematopoietic cells. *Development* **129**, 2003–2013 (2002).
16. Kovalenko, A. *et al.* Caspase-8 deficiency in epidermal keratinocytes triggers an inflammatory skin disease. *J. Exp. Med.* **206**, 2161–2177 (2009).
17. Lee, P. *et al.* Dynamic expression of epidermal caspase 8 simulates a wound healing response. *Nature* **458**, 519–523 (2009).
18. Kaufmann, T. *et al.* Fatal hepatitis mediated by tumor necrosis factor TNF α requires caspase-8 and involves the BH3-only proteins Bid and Bim. *Immunity* **30**, 56–66 (2009).
19. Hoebe, K. *et al.* Identification of *Lps2* as a key transducer of MyD88-independent TIR signalling. *Nature* **424**, 743–748 (2003).
20. Newton, K., Sun, X. & Dixit, V. M. Kinase RIP3 is dispensable for normal NF- κ Bs, signaling by the B-cell and T-cell receptors, tumor necrosis factor receptor 1, and Toll-like receptors 2 and 4. *Mol. Cell. Biol.* **24**, 1464–1469 (2004).
21. Beisner, D. R., Ch'en, I. L., Kolla, R. V., Hoffmann, A. & Hedrick, S. M. Cutting edge: innate immunity conferred by B cells is regulated by caspase-8. *J. Immunol.* **175**, 3469–3473 (2005).
22. Salmena, L. & Hakem, R. Caspase-8 deficiency in T cells leads to a lethal lymphoinfiltrative immune disorder. *J. Exp. Med.* **202**, 727–732 (2005).
23. Bidere, N., Su, H. C. & Lenardo, M. J. Genetic disorders of programmed cell death in the immune system. *Annu. Rev. Immunol.* **24**, 321–352 (2006).
24. Ch'en, I. L. *et al.* Antigen-mediated T cell expansion regulated by parallel pathways of death. *Proc. Natl Acad. Sci. USA* **105**, 17463–17468 (2008).
25. Laouar, Y. & Ezine, S. *In vivo* CD4+ lymph node T cells from lpr mice generate CD4-CD8-B220+ TCR-beta low cells. *J. Immunol.* **153**, 3948–3955 (1994).
26. Ventura, A. *et al.* Restoration of p53 function leads to tumour regression *in vivo*. *Nature* **445**, 661–665 (2007).
27. Kaiser, W. J. & Offermann, M. K. Apoptosis induced by the toll-like receptor adaptor TRIF is dependent on its receptor interacting protein homotypic interaction motif. *J. Immunol.* **174**, 4942–4952 (2005).
28. Belo, J. A. *et al.* Cerberus-like is a secreted factor with neutralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. *Mech. Dev.* **68**, 45–57 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge D. Green, A. Oberst and C. Dillon for their discussions and material support. We thank V. Dixit and K. Newton for *RIP3*^{-/-} mice, A. Kowalczyk, R. Oas, S. Speck and R. Ahmed for providing reagents, T. Kaufmann for technical advice, C. Strauss for editing, and A. L. McCormick for discussion and comments on the manuscript. Additionally, we appreciate M. Dowdy and J. Perry for mouse colony maintenance. This work was supported by the Georgia Cancer Coalition and the NIH (PHS grants R01 AI20211 and AI30363 to E.S.M.).

Author Contributions W.J.K., J.W.U., L.P.D.-B., and D.L.-R. designed and performed experiments and assembled figure panels. A.B.L. and T.C. guided evaluation of embryos. R.H. provided essential material support. E.S.M. supervised the project. W.J.K. and E.S.M. wrote the manuscript. T.C., R.H., J.W.U., L.P.D.-B., and D.L.-R. edited the text and figures during assembly of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to W.J.K. (wkaiser@emory.edu).

METHODS

Mice. *RIP3*^{-/-} mice have been described previously²⁰. TRIF mutant (strain C57BL/6J-Ticam1^{Lps2})¹⁹ mice were from Jackson Laboratory. *Casp8*^{+/-} mice were generated by crossing *Casp8*^{fl3-4/wt} (ref. 14) with Rosa-CreER mice²⁶. *Rip3*^{-/-} and *Casp8*^{+/-} were subsequently intercrossed. PCR genotyping of *Casp8*^{-/-} mice was performed with primers 5'-TTGAGAACAAGACCTGGGGACTG and 5'-GGATGTCAGGAAAAGATTGTGTGTC. PCR amplification allele produces a 750-bp band (wild-type *Casp8*), or a 200-bp band (*Casp8*^{-/-} allele produces). Genotyping of *Rip3*^{-/-} mice was performed with the primers 5'-CGCTTTAGAAGCCTCAGGTTGAC, 5'-GCAGGCTCTGGTGACAAGATTTCATGG, and 5'-CCAGAGGCCACTTGTGTAGCG. PCR produces a 700-bp band (wild-type *Rip3* allele) or a 450-bp band (*Rip3* deletion allele). Mice were bred and maintained by Emory University Division of Animal Resources where all procedures were approved by the Emory University Institutional Animal Care and Use Committee.

Cell culture, plasmids, transfections and transductions. L929 were maintained in DMEM containing 4.5 g ml⁻¹ glucose, 10% FBS (Atlanta Biologicals), 2 mM L-glutamine, 100 U ml⁻¹ penicillin and 100 U ml⁻¹ streptomycin (Invitrogen). For BMDM culture, pooled bone marrow cells from flushed tibias and femurs of indicated animals were differentiated for 5 to 7 days in DMEM containing 20% serum and 20% filtered L929 medium containing macrophage colony-stimulating factor (M-CSF). Cells were harvested with PBS containing 0.5 mM EDTA, seeded at a density of 3 × 10⁴ cells per well in a 96-well plate, and then cultured at least 18 h in DMEM containing 10% FBS before stimulation. Transient transfections were performed using Lipofectamine 2000 according to the manufacturer's protocol (Invitrogen). Plasmid encoding an amino-terminal Flag-epitope tagged M36 was generated by inserting M36 amplified from K181 strain MCMV genomic DNA into pQCXII (Clontech). The pLKO.1 based RIP3 (TRCN0000022535) shRNA (5'-CCGGCCTCAGATTCCACATACTTTACTCGAGTAAAGTATGTGGAATCTGAGGTTTTT) construct were obtained from Open Biosystems. The pLKO.1-Scramble control shRNA (5'-CCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGG GCGACTTAACCTTAGG) vector²⁹ and all other plasmids have been described^{6,27,30}. Lentiviral and retroviral production, infection, and selection of transduced cells have all been described previously³⁰. Inhibition of Caspase-8 expression in L929 cells employed murine Caspase-8 SMARTpool (L043044-00-0005) Dharmacon siRNAs consisting of 5'-GUGAAUGGAACCUGGUAUA, 5'-GUCACGACUUCAGACAAA, 5'-GAAGAUCCGAGGAUUAUGAA, and 5'-AGAGUUGUCUUUAUGCUAU, in comparison to OnTarget Plus Non-Targeting Pool (D-001810-10) siRNA Control. L929 cells seeded on 96-well plates were transfected with 4 pmol of siRNA with Lipofectamine 2000 according to the manufacturer's protocol.

Immunoblot and immunoprecipitations. Immunoblotting, preparation of protein extracts, and immunoprecipitations were as previously described²⁷. The following antibodies were used in IB analyses: mouse anti-β-actin (clone AC-74; Sigma), mouse anti-RIP1 (clone 38; BD Biosciences), rabbit anti-RIP3 (Imgenex), rat anti-Caspase-8 (clone 1G12; Axxora), anti-mouse IgG-HRP (Vector Laboratories), anti-rabbit IgG-HRP (Vector Laboratories), and anti-rat IgG-HRP (Jackson Laboratories).

Cell viability assays. Viability of L929 cells (5,000 cells per well) or BMDM (30,000 cells per well) seeded into 96-well plates was determined 96-h after transfection of siRNAs, 12 h after transfection of plasmids, or as indicated in the text. Where indicated, cells were stimulated in the presence of CHX (5 μg ml⁻¹) (Sigma). Nec-1 and z-VAD-fmk were from Calbiochem and Enzo Life Sciences, respectively. Cell viability was determined indirectly by measuring the intracellular levels of ATP using the Cell Titer-Glo Luminescent Cell Viability Assay kit (Promega) according to the manufacturer's instructions and was graphed relative to control cultures. Luminescence was measured on a Synergy HT Multi-Detection Microplate Reader (BioTek). To determine the number of surviving GFP positive cells, cells were transfected with pMaxGFP (Amara Biosystems) and analysed at 48 h post-transfection on an LSRII (BD Biosciences) using FlowJo software (Tree Star).

In situ hybridization. Whole-mount *in situ* hybridization was performed as described²⁸. Digoxigenin-labelled antisense RNA probes were transcribed from linearized RIP3 encoding plasmid (accession number BC029210, ATCC) according to the manufacturer's directions (Roche). For the hybridization results shown, the probe was synthesized from plasmid linearized with BamHI. A second RIP3 *in*

situ hybridization probe revealed a similar pattern of RIP3 expression (data not shown).

Flow cytometry. Single-cell suspensions were prepared from spleen, lymph nodes and thymus by mechanical disruption through a metal strainer. Bone marrow cells were isolated by flushing femurs with RPMI supplemented with 10% FBS, penicillin, streptomycin and 50 μM β-mercaptoethanol. Red blood cells were lysed using ammonium chloride solution (0.15 M NH₄Cl, 10 mM NaHCO₃, and 1.0 mM Na₂EDTA in H₂O, pH 7.4). Cells were resuspended in FACS staining buffer (PBS containing 0.2% BSA and 0.09% sodium azide), filtered through 40-μm screens and viable cells were enumerated using trypan blue exclusion. In all instances 1 × 10⁶ cells were prepared for flow cytometric analysis of surface antigens. Cells were incubated with 10% normal rat serum (Pel-Freez) and anti-mouse CD16/CD32 (2.4G2; BD Pharmingen) to reduce non-specific antibody interactions before incubating with lineage specific antibodies. For T-cell stimulation with anti-CD3 and anti-CD28 (BD Biosciences), splenocytes were incubated for 5 h in the presence of GolgiPlug (BD Biosciences) and subsequently evaluated for intracellular cytokine production using the Cytofix/Cytoperm kit (BD Biosciences) according to manufacturer's instructions and staining for intracellular cytokines IFNγ-FITC (clone XMG1.2 from BD) and TNFα-PE-Cy7 (clone MP6-XT22 from BD). The antibodies used were: Ly6C FITC (AL-21), B220 FITC (RA3-6B2), IgD PE (11-26c), CD19 PerCP-Cy5.5 (1D3), CD25 PerCP-Cy5.5 (PC61.5), CD11b APC-Cy7 (M1/70) and CD3 Pacific Blue (500A2), purchased from BD Biosciences; CD4 PE (GK1.5) and CD49b APC (DX5) purchased from eBioscience; CD62L PE-Cy7 (MEL-14) CD8 APC (53-6.7), and CD44 APC-Cy7 (IM7) purchased from BioLegend; and CD45 PE-Texas Red (30-F11) and F4/80 Pacific Orange (BM8) purchased from Invitrogen. Data were acquired using an LSRII flow cytometer (BD Biosciences) and analysed with FlowJo software. All statistical analyses were unpaired Student's *t*-tests using Prism (GraphPad Software).

Immunofluorescence microscopy. Yolk sacs were harvested and fixed for 2 h using 4% paraformaldehyde (Electron Microscopy Sciences) in PBS. Cells were permeabilized for 45 min in 0.25% Triton X-100 (Sigma) in PBS, blocked for 1 h in PBS containing 2% goat serum, and incubated 2 h at room temperature with rabbit polyclonal RIP3 antibody (Imgenex) along with PECAM Rat IgG2a (clone MEC 13.3; BD Biosciences) or FITC-conjugated CD41 Rat IgG1a (clone MwReg30; BD Biosciences) diluted in blocking buffer. Yolk sacs were washed multiple times with PBS and then incubated for 1 h at room temperature with Alexa Fluor 488 (or 594) goat anti-rabbit/rat antibody and Alexa Fluor 594 (or 488) goat anti-rabbit/rat antibody (Invitrogen). 4',6-diamidino-2-phenylindole (DAPI, Sigma) was used as a nuclear (DNA) counter stain. Following multiple washes with PBS, cells were mounted with Gel Mount (Biomedex). Images were acquired on a Carl-Zeiss LSM 510 META confocal fluorescence microscope.

DR-induced hepatitis. Fatal hepatitis was induced by intraperitoneal injection of with 12.5 μg of anti-mouse Fas antibody (clone Jo2, BD Biosciences) or for LPS+GalN induced hepatitis, mice were injected intraperitoneally with 100 ng of ultrapure LPS K12 (InvivoGen) in the presence of 20 mg of the liver-specific transcriptional inhibitor D-(+)-galactosamine (GalN, SIGMA). The organs evaluated for histology were fixed in 10% neutral buffered formalin before paraffin embedding, processing and staining for haematoxylin and eosin. Tissue processing and staining was performed by Emory University's Division of Animal Resources (DAR).

MCMV Infections. Mice were inoculated by i.p. (intraperitoneal) injection with 10⁶ plaque-forming units (p.f.u.) salivary-gland-derived v70 strain of MCMV (a gift of C. Biron). Salivary gland viral stock was generated by sonicating submaxillary glands of 8-week-old BALB/c mice infected with 1 × 10³ p.f.u. of v70 strain of MCMV at 14 days post infection.

HPC progenitor assay. For yolk sac-derived clonogenic progenitor assays, single cell suspensions were prepared from E10.5/E11.5 yolk sacs in 2 ml of medium (M3434; StemCell Technologies). Cells were grown for 12 days at 37 °C in 5% CO₂ in the absence or presence of 30 μM Nec-1. Colonies were scored by microscopic analysis.

29. Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098–1101 (2005).

30. Kaiser, W. J., Upton, J. W. & Mocarski, E. S. Receptor-interacting protein homotypic interaction motif-dependent control of NF-κB activation via the DNA-dependent activator of IFN regulatory factors. *J. Immunol.* **181**, 6427–6434 (2008).

Functional complementation between FADD and RIP1 in embryos and lymphocytes

Haibing Zhang¹, Xiaohui Zhou^{1†}, Thomas McQuade², Jinghe Li¹, Francis Ka-Ming Chan² & Jianke Zhang¹

FADD is a common adaptor shared by several death receptors for signalling apoptosis through recruitment and activation of caspase 8 (refs 1–3). Death receptors are essential for immune homeostasis, but dispensable during embryogenesis. Surprisingly, *Fadd*^{−/−} mice die *in utero*^{4,5} and conditional deletion of FADD leads to impaired lymphocyte proliferation^{6,7}. How FADD regulates embryogenesis and lymphocyte responses has been a long-standing enigma. FADD could directly bind to RIP1 (also known as RIPK1), a serine/threonine kinase that mediates both necrosis and NF-κB activation. Here we show that *Fadd*^{−/−} embryos contain raised levels of RIP1 and exhibit massive necrosis. To investigate a potential *in vivo* functional interaction between RIP1 and FADD, null alleles of RIP1 were crossed into *Fadd*^{−/−} mice. Notably, RIP1 deficiency allowed normal embryogenesis of *Fadd*^{−/−} mice. Conversely, the developmental defect of *Rip1*^{−/−} lymphocytes was partially corrected by FADD deletion. Furthermore, RIP1 deficiency fully restored normal proliferation in *Fadd*^{−/−} T cells but not in *Fadd*^{−/−} B cells. *Fadd*^{−/−}*Rip1*^{−/−} double-knockout T cells are resistant to death induced by Fas or TNF-α and show reduced NF-κB activity. Therefore, our data demonstrate an unexpected cell-type-specific interplay between FADD and RIP1, which is critical for the regulation of apoptosis and necrosis during embryogenesis and lymphocyte function.

Programmed cell death—including apoptosis and necrosis—is a fundamental biological process that is essential during embryonic development and for homeostasis in somatic tissues. Death receptors can signal apoptotic cell death when engaged by their cognate ligands⁸. This extrinsic death pathway requires the adaptor protein FADD, which couples the signal generated by death receptors to the apical caspase 8 (refs 3, 9, 10). The resulting activation of caspase 8 triggers a battery of downstream caspases, leading to apoptosis. Recently, death receptors were shown to induce necrosis-like cell death in the presence of caspase inhibitors^{11,12}. Death-receptor-induced necrosis is blocked in cells lacking the protein serine/threonine kinases RIP1 and RIP3^{11–15}. FADD and RIP1 have indispensable roles in development, as *Fadd*^{−/−} mice die during mid-gestation stages^{4,5}, and *Rip1*^{−/−} mice die at birth¹⁶. Whereas the developmental defect in *Rip1*^{−/−} mice is presumably due in part to defective NF-κB activation and increased cell death¹⁷, the mechanism that underlies the developmental defect of *Fadd*^{−/−} mice has remained elusive.

Developmental retardation of *Fadd*^{−/−} mouse is apparent at embryonic day (E)11.5 to 13.5 (Supplementary Fig. 1a, b). Histological analysis revealed extensive necrotic cell death and cell loss in *Fadd*^{−/−} embryos (Fig. 1a). Absence of FADD sensitizes human Jurkat T lymphoma cells to tumour necrosis factor (TNF)-induced necrosis^{11,12}. Furthermore, *Fadd*^{−/−} mouse embryonic fibroblasts (MEFs) were hypersensitive to reactive oxygen species (ROS)-induced necrosis¹⁸. In contrast, necrosis induced by ROS was blocked in *Rip1*^{−/−} MEF cells. Examination of wild-type E12.5 embryos showed that RIP1 was expressed at low levels whereas RIP3 was readily detected in multiple tissues including the nervous system, heart and lung (Fig. 1a and

Supplementary Fig. 2). Interestingly, RIP1 expression is highly elevated in *Fadd*^{−/−} embryos (Fig. 1a). Furthermore, punctate RIP3 staining was observed in cells of *Fadd*^{−/−} embryos, which is indicative of RIP3 aggregation and activation¹⁴. These results indicate that induction of RIP1 expression might have a role in the necrosis observed in *Fadd*^{−/−} embryos.

To investigate a potential *in vivo* molecular interplay between FADD and RIP1-mediated signalling, we crossed the *Rip1* knockout alleles into *Fadd*^{−/−} mice. Notably, *Fadd*^{−/−}*Rip1*^{−/−} double-knockout (DKO) embryos were detected at E14.5 at the expected Mendelian frequencies (Table 1, Fig. 1b and Supplementary Fig. 1c–f). In sharp contrast to the highly deformed E14.5 *Fadd*^{−/−} embryos, E14.5 DKO embryos were indistinguishable from wild-type control embryos (Fig. 1b). DKO embryos of normal morphology were also found at later gestation stages E15.5, E16.5, E17.5 and E18.5 at the expected Mendelian frequencies (Supplementary Fig. 1c and Table 1). Importantly, live DKO neonates were also detected at birth (Supplementary Fig. 1d). No *Fadd*^{−/−} embryos were detected at E15.5 or later stages. Histological examination did not reveal obvious defects in *Rip1*^{−/−} and DKO E18.5 embryos

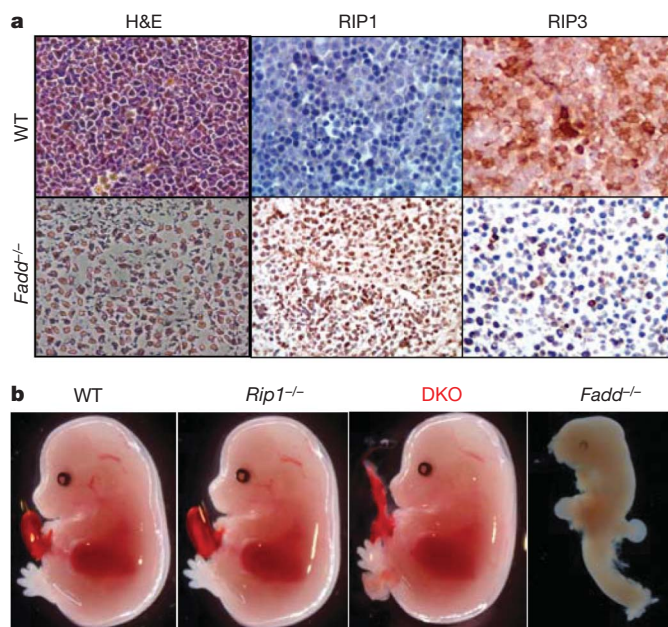


Figure 1 | RIP1 deficiency rescues *Fadd*^{−/−} mice from embryonic necrosis and lethality. **a**, *Fadd*^{−/−} embryos exhibit massive necrosis and altered RIP1 and RIP3 expression. E12.5 wild-type (WT, top panels) or *Fadd*^{−/−} embryos (bottom panels) were fixed in formalin. Left panels of haematoxylin and eosin (H&E) staining show extensive cell loss and pyknotic nuclei in the *Fadd*^{−/−} fetal liver. Middle and right panels indicate staining for RIP1 or RIP3 (magnification ×600). **b**, E14.5 DKO embryos appear normal, contrasting with the defective *Fadd*^{−/−} embryos.

¹Department of Microbiology and Immunology, Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, USA. ²Department of Pathology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA. †Present address: College of Life Science, Wenzhou Medical College, Zhejiang 325035, China.

Table 1 | Genetic analysis of *Fadd* and *Rip1* deficiency in mice

Genotype	<i>Fadd</i> <i>Rip1</i>	+/+ or +/- +/+ or +/-	+/+ or +/- -/-	-/- +/+ or +/-	-/- -/-	Total
E14.5–18.5 embryos	Actual	102	45	17	14*	178
	Expected	100	33	33	11	
At or after birth						
Day 0	Actual	26	4	0	1*	31
	Expected	17	6	6	2	
>Day 0	Actual	79	19	0	6*	104
	Expected	59	20	20	7	

Timed pregnancy analyses were set up to determine the frequencies of each genotype of offspring from intercrosses of *Fadd*^{-/-}*Rip1*^{-/-} mice. The actual and expected numbers of each or groups of genotypes were shown. Neonates were killed and analysed at birth (Day 0) or monitored for survival for up to 6 months (>Day 0). Asterisks indicate numbers of DKO embryos or neonates. Postnatal *Rip1*^{-/-} and DKO neonates (>Day 0) died within 4 days after birth.

(Supplementary Fig. 3). Postnatal monitoring was performed to determine the survival of DKO mice. Among the 104 postnatal mice analysed (>Day 0, Table 1), 30 died within 4 days after birth, and this group contained 6 DKO and 19 *Rip1*^{-/-} genotypes. No death was observed after postnatal day 4 and DKO mice were not present in the remaining 74 mice that survived beyond 3 weeks (Table 1). These results demonstrate that RIP1 deficiency fully restores embryonic development of *Fadd*^{-/-} mice. However, loss of FADD does not prevent neonatal lethality of *Rip1*^{-/-} mice. ROS is an important effector mechanism for necrotic cell death. *Fadd*^{-/-} MEFs were hypersensitive to ROS-induced death (Supplementary Fig. 4a, b). In contrast, *Rip1*^{-/-} MEFs were highly resistant to ROS-induced necrosis¹⁸. Addition of the RIP1-specific inhibitor necrostatin 1 (Nec1)¹⁹ greatly reduced ROS hypersensitivity in *Fadd*^{-/-} MEFs (Supplementary Fig. 4a, b). Importantly, DKO MEFs were resistant to ROS-induced death. Collectively, these results indicate that FADD deficiency primes embryonic cells to ROS- and RIP1-dependent necrosis, which might cause embryonic lethality.

Although important at early haematopoietic stages²⁰, FADD has a minor role in post-lineage commitment lymphopoiesis^{4,6,7}. Although *Rip1*^{-/-} neonates contain normal thymocyte numbers¹⁶, *Rip1*^{-/-} fetal liver cells failed to reconstitute the peripheral T-cell compartment¹⁷. The rescue of embryonic development in DKO mice prompted us to examine whether a similar FADD–RIP1 interaction might regulate lymphocyte development. To this end, we adoptively transferred fetal liver cells containing haematopoietic progenitor cells into immunodeficient NSG recipient mice. In agreement with previous results¹⁷, NSG chimaeras reconstituted with *Rip1*^{-/-} fetal liver cells contained markedly reduced CD4⁺CD8⁺ double-positive immature and CD4⁺ or CD8⁺ single-positive mature thymocytes (Fig. 2a). In contrast, the thymic population profile of DKO fetal liver cell chimaeras was similar to that of the wild-type control thymus (Fig. 2a). Reconstitution of the peripheral lymphoid compartment by DKO fetal liver cells was apparent, as indicated by the spleen sizes of the recipients of DKO fetal liver cells, which were similar to the size of the control spleens receiving wild-type fetal liver cells (Supplementary Fig. 5a). In contrast, the spleen of *Rip1*^{-/-} chimaeras was smaller than that of wild-type or DKO chimaeras. Flow cytometric analyses showed that *Rip1*^{-/-} chimaeras contained few CD3⁺ T cells in the periphery (Fig. 2b and Supplementary Fig. 5b, c). In contrast, DKO chimaeras contained significantly higher numbers of T cells in the spleen, lymph nodes and blood (Fig. 2b and Supplementary Fig. 5b, c). Similarly, FADD deficiency partially rescued *Rip1*^{-/-} B-cell development (Supplementary Fig. 5c). *Rip1*^{-/-} thymocytes were readily killed by treatments with anti-Fas antibodies or TNF (Fig. 2c, d). Interestingly, DKO thymocytes were highly resistant to these death stimuli. Although FADD deficiency fully reversed the hypersensitivity to Fas- and TNF- α -induced killing, it only partially corrected the NF- κ B activation defect in *Rip1*^{-/-} T cells, B cells and MEFs (Fig. 2e–g). These results indicate that the partial rescue of lymphocyte development in the DKO chimaeras is due to inhibition of FADD-mediated apoptosis rather than rescue of NF- κ B activation.

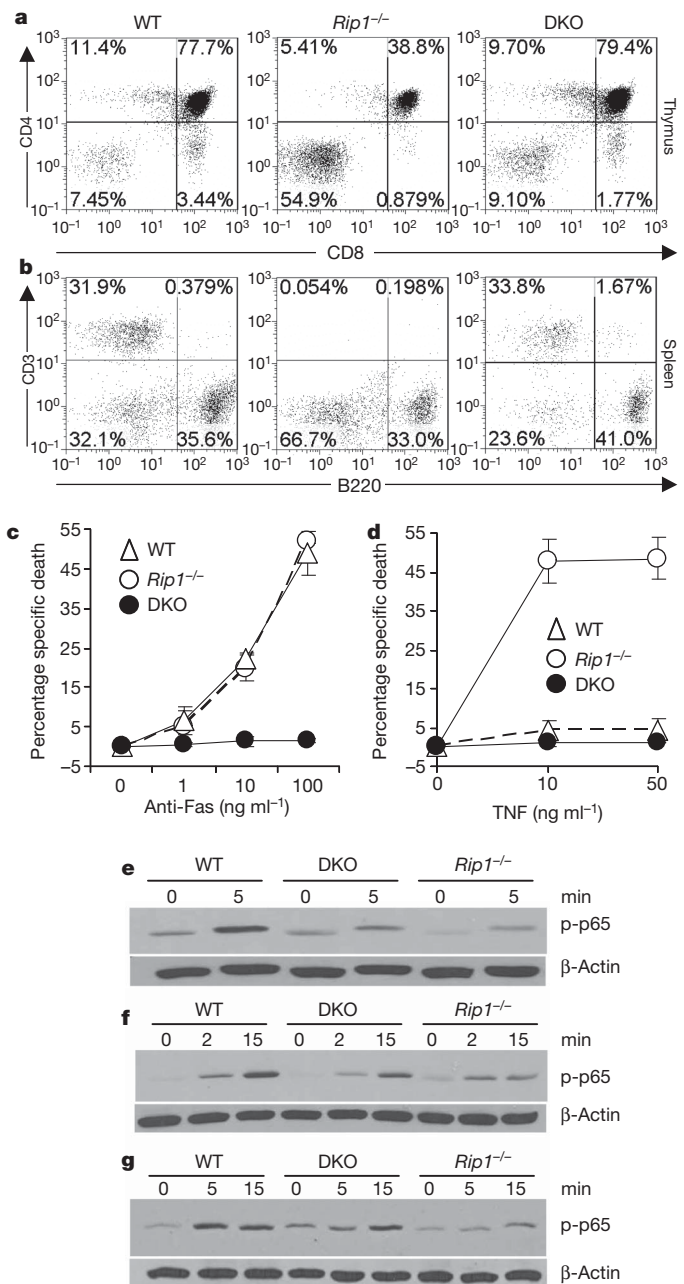


Figure 2 | FADD deficiency partially corrects the *Rip1*^{-/-} T-cell developmental defect by blocking apoptosis. **a, b**, Lymphocytes in the chimaeras of the indicated genotypes were analysed. **c, d**, E18.5 fetal thymocytes of the indicated genotypes were treated with anti-Fas antibodies (**c**) or TNF- α (**d**) and death responses determined 12 h after stimulation. Error bars represent mean \pm s.e.m. of triplicates. **e–g**, T cells (**e**) and B cells (**f**) from NSG chimaeras of the indicated genotypes were stimulated with anti-CD3 and anti-CD28 antibodies or with LPS, respectively. MEFs were stimulated with TNF- α (**g**). NF- κ B activation is indicated by the induction of p65 phosphorylation (p-p65). β -Actin, loading controls.

Although T-cell-specific deletion of FADD had no effect on thymic development, the resulting mature *Fadd*^{-/-} T cells were highly defective in T-cell antigen receptor (TCR)-induced proliferation⁶ (Fig. 3a). When compared to *Fadd*^{-/-} and wild-type controls, DKO T cells stimulated through the TCR/CD28 showed a remarkable rescue in their proliferative responses (Fig. 3a and Supplementary Table 1a). When transferred to hosts with the genes encoding TCR- α and TCR- β deleted (*Tcr α /b*^{-/-}), DKO T cells were functionally competent to expand and produce IFN- γ in response to challenge with Pichinde virus (PV) (Fig. 3b). Acute

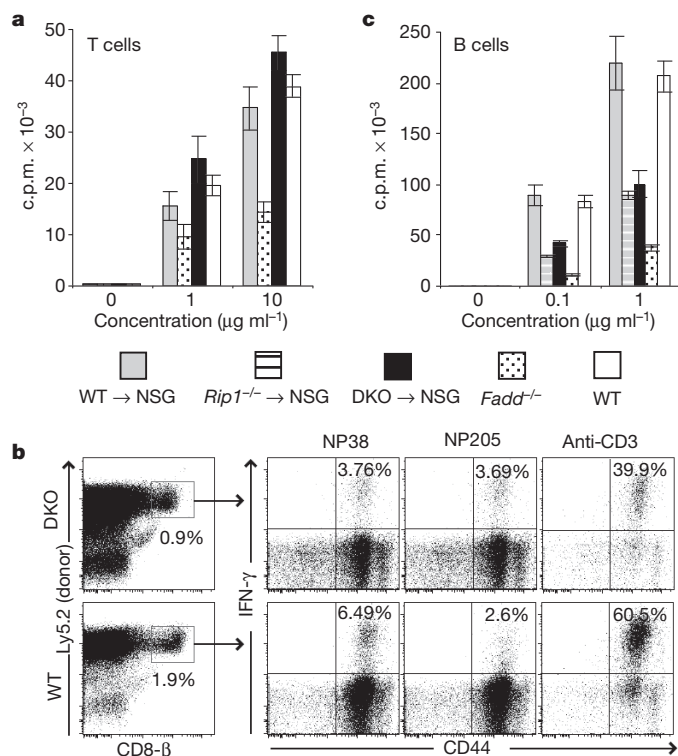


Figure 3 | RIP1 deficiency rescues the *Fadd*^{-/-} T-cell proliferation defect. **a**, T cells of the indicated wild-type and chimaera genotypes were stimulated with anti-CD3 and anti-CD28 antibodies. Proliferation was measured by [³H]-thymidine incorporation (c.p.m., counts per minute). **b**, Splenocytes of the indicated genotypes were transferred into *Tcrb*^{-/-} hosts and challenged with PV. Acute responses of the donor CD8 T cells to the indicated epitopes were determined. Anti-CD3 antibodies were used as control. **c**, B cells of the indicated genotypes were stimulated with LPS and proliferation was measured as in **a**. Results shown (**a** and **c**) are mean ± s.e.m. of triplicates.

CD8⁺ T-cell responses to the immunodominant epitope NP38 and subdominant epitope NP205 were similar between wild-type and DKO donor cells (Fig. 3b). Moreover, challenge of wild-type hosts adoptively transferred with DKO lymphocytes with lymphocytic choriomeningitis virus (LCMV) showed that the DKO T cells could generate a productive antiviral response to the immunodominant epitope NP396 (Supplementary Fig. 6a). Collectively, these results indicate that RIP1-dependent necrosis underlies the defective proliferation in *Fadd*^{-/-} T cells and that inactivation of RIP1 restores the proliferative capacity of *Fadd*^{-/-} T cells.

Although FADD does not have a significant role in B-cell antigen receptor (BCR)- or CD40-induced proliferation responses, it is required for TLR3 and TLR4 signalling in B cells⁷. Consistent with these observations, wild-type, *Rip1*^{-/-}, *Fadd*^{-/-} and DKO B cells responded similarly to stimulation with anti-IgM or anti-CD40 antibodies (Supplementary Fig. 6b and data not shown). In contrast to the rescue of T-cell proliferation, DKO B cells remained defective in proliferative responses to the TLR3 and TLR4 agonists poly(I:C) and lipopolysaccharide (LPS), respectively (Fig. 3c, Supplementary Fig. 6c and Supplementary Table 1b). The differential effect of RIP1 deletion on *Fadd*^{-/-} T- and B-cell proliferation was confirmed with the RIP1-specific inhibitor Nec1 (Supplementary Fig. 6d, e). Although caspase inhibition did hamper LPS-induced B-cell proliferation (Supplementary Fig. 7a, b), RIP1 cleavage, which inactivates the pro-necrotic activity of RIP1, was only observed in TCR-treated T cells but not LPS-induced B cells or ROS-treated MEFs (Supplementary Fig. 8a–c). Therefore, the FADD–RIP1 axis is preferentially required for controlling proliferation in T cells, but not in B cells.

This study demonstrates that absence of RIP1 restores normal embryogenesis of *Fadd*^{-/-} mice, and that FADD deficiency partially corrects the developmental defect in *Rip1*^{-/-} T cells. This finding provides compelling genetic evidence that a critical *in vivo* role for FADD during embryogenesis is to inhibit RIP1-mediated necrosis. In T cells, RIP1 is required to help suppress FADD-mediated apoptosis. Interestingly, caspase 8^{-/-} mice showed embryonic and T-cell defects similar to that of *Fadd*^{-/-} mice^{4,5,21}. Moreover, Nec1 rescued the proliferative defect of caspase 8^{-/-} and *Fadd*^{-/-} T cells^{22,23}. Hence, FADD and caspase 8 probably act in concert to keep RIP1-mediated necrosis in check by cleavage and inactivation of RIP1. Such a regulatory mechanism is crucial to ensure proper embryogenesis and to prevent abortive expansion of T cells during immune responses. Our results also reveal an unexpected role for RIP1 in keeping FADD-mediated apoptosis at bay during T-cell development. This regulation is not entirely dependent on RIP1-mediated activation of the NF-κB pathways, unlike that proposed previously^{16,17}. The reduced NF-κB activation had no major effect on DKO T-cell proliferations. In addition to lymphocytes, impairment of the NF-κB pathway is also present in *Rip1*^{-/-} and DKO MEFs and probably other cell types, which might lead to postnatal lethality as seen in *Rip1*^{-/-} and DKO mice. In summary, our results reveal a complex functional interaction between FADD and RIP1 that is context and cell-type dependent.

METHODS SUMMARY

Heterozygous *Fadd*^{+/-} mutant mice have previously been described⁴. Heterozygous *Rip1*^{+/-} mutant mice¹⁶ were provided by M. Kelliher and crossed into *Fadd*^{+/-} mice. Timed pregnancy was set up with the resulting *Fadd*^{+/-} *Rip1*^{+/-} mouse intercrosses. Embryos were isolated and genotyped by PCR using tissue genomic DNA templates and confirmed by western blots. Fetal liver cells were isolated from E14.5 embryos and adoptively transferred into irradiated (200 rad) NSG (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl/SzJ}) mice purchased from Jackson Laboratory. All animal studies were approved by Institutional Animal Care and Use Committees. Ten-to-twelve weeks after transfer, cells were isolated from the indicated organs and were analysed by flow cytometry. T cells and B cells were purified from the spleen and lymph nodes by high-speed sorting and proliferative responses were analysed as described previously^{6,7}. *Fadd*^{-/-} mutant T cells and B cells were isolated by sorting from T-cell-specific and B-cell-specific *Fadd*^{-/-} mice as described previously^{6,7}. Western blotting was performed according to standard protocols. For embryonic cell death assays, primary MEFs were prepared following the NIH 3T3 protocol. *Fadd*^{-/-} MEFs were prepared from E8.5 embryos and DKO MEFs from E14.5 MEFs. MEF cells were cultured in complete DMEM to 80% confluence, and treated with 0.5 mM H₂O₂ with or without Nec1 (50 µM) for 12 h, and cell death was determined by propidium iodide staining and flow cytometry. Images were taken by using a Nikon inverted light microscope. For virus infections, after adoptive transfer of lymphocytes, mice were challenged with 5 × 10⁴ p.f.u. of LCMV or 1 × 10⁷ p.f.u. of PV. Peptide-specific CD8 T-cell responses were measured 8 days after infection by intracellular IFN-γ staining.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 December 2010; accepted 27 January 2011.

Published online 2 March 2011.

1. Boldin, M. P. *et al.* A novel protein that interacts with the death domain of Fas/APO1 contains a sequence motif related to the death domain. *J. Biol. Chem.* **270**, 7795–7798 (1995).
2. Chinnaiyan, A. M., O'Rourke, K., Tewari, M. & Dixit, V. M. FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis. *Cell* **81**, 505–512 (1995).
3. Zhang, J. & Winoto, A. A mouse Fas-associated protein with homology to the human Mort1/FADD protein is essential for Fas-induced apoptosis. *Mol. Cell. Biol.* **16**, 2756–2763 (1996).
4. Zhang, J., Cado, D., Chen, A., Kabra, N. H. & Winoto, A. Fas-mediated apoptosis and activation-induced T-cell proliferation are defective in mice lacking FADD/Mort1. *Nature* **392**, 296–300 (1998).
5. Yeh, W.-C. *et al.* FADD: essential for embryo development and signaling from some, but not all, inducers of apoptosis. *Science* **279**, 1954–1958 (1998).
6. Zhang, Y. *et al.* Conditional Fas-associated death domain protein (FADD):GFP knockout mice reveal FADD is dispensable in thymic development but essential in peripheral T cell homeostasis. *J. Immunol.* **175**, 3033–3044 (2005).

7. Imtiyaz, H. Z. *et al.* The Fas-associated death domain protein is required in apoptosis and TLR-induced proliferative responses in B cells. *J. Immunol.* **176**, 6852–6861 (2006).
8. Nagata, S. Apoptosis by death factor. *Cell* **88**, 355–365 (1997).
9. Boldin, M. P., Goncharov, T. M., Goltsev, Y. V. & Wallach, D. Involvement of MACH, a novel MORT1/FADD-interacting protease, in Fas/APO-1- and TNF receptor-induced cell death. *Cell* **85**, 803–815 (1996).
10. Muzio, M. *et al.* FLICE, a novel FADD-homologous ICE/CED-3-like protease, is recruited to the CD95 (Fas/APO-1) death-inducing signaling complex. *Cell* **85**, 817–827 (1996).
11. Holler, N. *et al.* Fas triggers an alternative, caspase-8-independent cell death pathway using the kinase RIP as effector molecule. *Nature Immunol.* **1**, 489–495 (2000).
12. Chan, F. K.-M. *et al.* A role for tumor necrosis factor receptor-2 and receptor-interacting protein in programmed necrosis and antiviral responses. *J. Biol. Chem.* **278**, 51613–51621 (2003).
13. Cho, Y. S. *et al.* Phosphorylation-driven assembly of the RIP1–RIP3 complex regulates programmed necrosis and virus-induced inflammation. *Cell* **137**, 1112–1123 (2009).
14. He, S. *et al.* Receptor interacting protein kinase-3 determines cellular necrotic response to TNF- α . *Cell* **137**, 1100–1111 (2009).
15. Zhang, D. W. *et al.* RIP3, an energy metabolism regulator that switches TNF-induced cell death from apoptosis to necrosis. *Science* **325**, 332–336 (2009).
16. Kelliher, M. A. *et al.* The death domain kinase RIP mediates the TNF-induced NF- κ B signal. *Immunity* **8**, 297–303 (1998).
17. Cusson, N., Oikemus, S., Kilpatrick, E. D., Cunningham, L. & Kelliher, M. The death domain kinase RIP protects thymocytes from tumor necrosis factor receptor type 2-induced cell death. *J. Exp. Med.* **196**, 15–26 (2002).
18. Shen, H. M. *et al.* Essential roles of receptor-interacting protein and TRAF2 in oxidative stress-induced cell death. *Mol. Cell. Biol.* **24**, 5914–5922 (2004).
19. Degterev, A. *et al.* Chemical inhibitor of nonapoptotic cell death with therapeutic potential for ischemic brain injury. *Nature Chem. Biol.* **1**, 112–119 (2005).
20. Rosenberg, S., Zhang, H. & Zhang, J. FADD deficiency impairs early hematopoiesis in the bone marrow. *J. Immunol.* **186**, 203–213 (2011).
21. Varfolomeev, E. E. *et al.* Targeted disruption of the mouse caspase 8 gene ablates cell death induction by the TNF receptors, Fas/Apo1, and DR3 and is lethal prenatally. *Immunity* **9**, 267–276 (1998).
22. Ch'en, I. L. *et al.* Antigen-mediated T cell expansion regulated by parallel pathways of death. *Proc. Natl Acad. Sci. USA* **105**, 17463–17468 (2008).
23. Osborn, S. L. *et al.* Fas-associated death domain (FADD) is a negative regulator of T-cell receptor-mediated necroptosis. *Proc. Natl Acad. Sci. USA* **107**, 13034–13039 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Kelliher for providing *Rip1^{+/-}* mice, S. Rosenberg, S. Waggoner, R. Welsh, V. Vanguri and Y. Liu for advice and technical assistance, X. Lin for discussions and suggestions, C. E. Calkins and K. Reinersmann for critical reading of the manuscript, and Z. Zhong for help with histology analysis. This study was supported in part by NIH grants CA95454, AI083915 and AI076788 awarded to J.Z. and AI083497 awarded to F.K.-M.C.; a W. W. Smith Charitable Trust grant, a TJU Enhancement grant, and a CONCERN Foundation grant awarded to J.Z. F.K.-M.C. is a member of the UMass DERC (DK32520) and is also supported in part by an NIH grant (AI017672).

Author Contributions J.Z. conceived and initiated the project. J.Z., F.K.-M.C. and H.Z. planned the experiments. All authors performed and analysed the experiments. J.Z. and F.K.-M.C. wrote the manuscript with contribution from H.Z.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.Z. (jzhang@mail.jci.tju.edu) or F.K.-M.C. (francis.chan@umassmed.edu).

METHODS

Reagents. Antibodies were purchased from the following vendors: RIP1 (BD Pharmingen), RIP3 (ProSci), caspase 8 (Axxora), phospho-p65 (Cell Signaling), CD3 (BD Biosciences), CD28 (BD Biosciences), and streptavidin (BD Biosciences). Antibodies used in immunophenotyping were purchased from BD Biosciences or eBioscience or BioLegend. Rabbit anti-FADD antibodies were generated in house. Necrostatin 1 and zVAD-fmk were purchased from Axxora.

Mouse genetic analysis. Mating of *Fadd*^{+/-}*Rip1*^{+/-} double heterozygous mice was set up in the evening. Female mice were examined in the early morning and embryos were designated E0.5 on the day the vaginal plug was detected. E11.5 to E18.5 embryos were isolated from pregnant mice and genomic DNA was extracted from embryo tissues upon digestion with protease K in an SDS-containing lysis buffer. Genotyping was performed by PCR using allele-specific primers. 5'-TGCG CCGACACGATCTACTG-3' and 5'-AGCTGTAGGCTTGTTCAGGGTGTTC-3' were used for the detection of the *Fadd* wild-type allele; 5'-ACTGTAGTGC CCAGCAGAGACCAGC-3' and 5'-CGTCTGGTGTTCGAGGCCACACGC-3' for the *Fadd* knockout allele; 5'-TGTGTCAAGTCTCCCTGCAG-3' and 5'-CAC GGTCTTTTGCCCTG-3' for the *Rip1* wild-type allele; and 5'-CTGCTA AAGCGCATGCTC-3' and 5'-CACGGTCTTTTGCCCTG-3' for the *Rip1* knockout allele. Presence and absence of the FADD and RIP1 proteins was confirmed by western blotting using anti-FADD and anti-RIP1 (BD Pharmingen) antibodies. For postnatal analysis, pregnant females were closely monitored around the time of delivery (after 18.5 days post coitum), by checking at least twice daily. Neonates were either killed at birth for genotyping or monitored for survival for up to 6 months. Dead mice were immediately removed for genotyping.

Histological analysis and immunohistochemistry. Embryos were fixed in 10% buffered formalin overnight. After dehydration in 70% and increasing concentrations of ethanol, embryos were embedded in paraffin. Embryo sections were prepared and standard H&E staining performed. Immunohistochemistry was performed using anti-RIP1 (H207, Santa Cruz) and anti-RIP3 (ProSci) antibodies at 1:200 dilution, followed by incubation with ImmPress anti-rabbit peroxidase (Vector Lab). Signals were developed with 3,3'-diaminobenzidine (5 min) and Enhancer (DAKO) for 2 min. Sections were then counterstained with haematoxylin.

Lymphocyte immunophenotyping. Lymphocytes were isolated from spleen, thymus, lymphocyte nodes and blood. Red blood cells were lysed using the ACK lysis buffer and single-cell suspension was prepared, stained for cell surface markers following standard protocols, and data acquisition was performed using a Coulter XL cytometer. Flow cytometric data were analysed with the FlowJo software (Treestar).

Cell death analysis. Thymocytes were isolated from the E18.5 embryos and resuspended in RPMI 1640 medium containing 10% fetal bovine serum and antibiotics. Cells were then seeded (1×10^5 per well) into 96-well flat-bottom plates with various

concentrations of anti-Fas antibodies (Jo2; BD Pharmingen) or recombinant mouse TNF- α (Biosource/Invitrogen) in the presence of $30 \mu\text{g ml}^{-1}$ cycloheximide (Sigma-Aldrich). Cell death was determined by propidium iodide (PI) staining and flow cytometry 12–16 h after stimulation.

Lymphocyte activation and western blot analysis. Peripheral T cells and B cells were purified from the spleen and lymphoid nodes using EasySep Isolation reagent (StemCell Technologies) or high-speed sorting. Purified T cells were stimulated with biotinylated anti-CD3 antibodies ($10 \mu\text{g ml}^{-1}$) and anti-CD28 antibodies ($2 \mu\text{g ml}^{-1}$) by crosslinking with streptavidin ($10 \mu\text{g ml}^{-1}$) for the indicated times at 37°C . In some experiments, cells were stimulated with plate-bound anti-CD3 and anti-CD28 antibodies. Purified B cells were stimulated with $10 \mu\text{g ml}^{-1}$ LPS for the indicated times at 37°C . Where indicated, cells were pre-incubated with $50 \mu\text{M}$ zVAD-fmk or $10 \mu\text{M}$ necrostatin 1 before stimulation. Cells were washed once in ice-cold PBS and lysed with ice-cold RIPA lysis buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% Nonidet P-40, 0.5% deoxycholate, 0.1% SDS, 1 mM phenylmethyl sulphonyl fluoride (PMSF) and a protease inhibitor cocktail (Roche). Protein was separated on 10% SDS-PAGE then transferred onto nitrocellulose membrane, blocked with 5% milk in TBST, incubated overnight at 4°C with primary antibodies. Blots were washed 3 times with TBST and incubated in HRP-conjugated second antibody as per the manufacturer's recommendations. The Western Lighting Chemiluminescence Reagent Plus (PerkinElmer) or ThermoScientific West Pico chemiluminescence reagents were used to detect signals on X-ray films (Kodak).

Cell culture and proliferation assays. Purified T cells and B cells were cultured in RPMI 1640 (Mediatech) supplemented with 100 U ml^{-1} penicillin, $100 \mu\text{g ml}^{-1}$ streptomycin, 1 mM glutamine, $5 \mu\text{M}$ β -mercaptoethanol (Sigma), and 10% FBS (HyClone Laboratories). Cells were cultured at 37°C in a 5% CO_2 incubator. B cells were cultured with the indicated concentrations of LPS, poly(I:C) (Sigma-Aldrich), goat anti-mouse IgM (F(ab')₂ μ chain, Jackson ImmunoResearch) for 48 h. T cells were stimulated with various concentration of plate-bound anti-CD3 antibodies in 96-well plates in the presence of anti-CD28 (200 ng ml^{-1}) antibodies for 48 h. Incorporated [³H]-thymidine was measured using a Wallac beta counter (PerkinElmer). To determine cell division, cell proliferation measured by dilution of CellTracer Violet dye (Invitrogen) was performed as per the manufacturer's instructions.

Virus infection. Ten-to-twenty million wild-type or DKO total splenocytes were injected via the tail vein into wild-type Ly5.1 congenic hosts. For injection into *Tcr α /b*^{-/-} hosts, 2×10^6 cells were injected. Eighteen hours later, mice were infected with 5×10^4 p.f.u. of lymphocytic choriomeningitis virus (Armstrong strain) or 1×10^7 p.f.u. of PV (strain AN3739) intraperitoneally. Eight days after infection, splenocytes were isolated and stimulated with the indicated viral peptide antigens for 5 h or anti-CD3 antibody for 4 h. Production of IFN- γ was examined by intracellular FACS staining on an LSRII flow cytometer (BD Biosciences).

MHC class II transactivator *CIITA* is a recurrent gene fusion partner in lymphoid cancers

Christian Steidl^{1*}, Sohrab P. Shah^{1*}, Bruce W. Woolcock¹, Lixin Rui², Masahiro Kawahara³, Pedro Farinha¹, Nathalie A. Johnson¹, Yongjun Zhao⁴, Adele Telenius¹, Susana Ben Neriah¹, Andrew McPherson¹, Barbara Meissner¹, Ujunwa C. Okoye³, Arjan Diepstra⁵, Anke van den Berg⁵, Mark Sun¹, Gillian Leung¹, Steven J. Jones⁴, Joseph M. Connors⁶, David G. Huntsman¹, Kerry J. Savage⁶, Lisa M. Rimsza⁷, Douglas E. Horsman¹, Louis M. Staudt², Ulrich Steidl³, Marco A. Marra^{4,8} & Randy D. Gascoyne¹

Chromosomal translocations are critically involved in the molecular pathogenesis of B-cell lymphomas, and highly recurrent and specific rearrangements have defined distinct molecular subtypes linked to unique clinicopathological features^{1,2}. In contrast, several well-characterized lymphoma entities still lack disease-defining translocation events. To identify novel fusion transcripts resulting from translocations, we investigated two Hodgkin lymphoma cell lines by whole-transcriptome paired-end sequencing (RNA-seq). Here we show a highly expressed gene fusion involving the major histocompatibility complex (MHC) class II transactivator *CIITA* (*MHC2TA*) in KM-H2 cells. In a subsequent evaluation of 263 B-cell lymphomas, we also demonstrate that genomic *CIITA* breaks are highly recurrent in primary mediastinal B-cell lymphoma (38%) and classical Hodgkin lymphoma (cHL) (15%). Furthermore, we find that *CIITA* is a promiscuous partner of various in-frame gene fusions, and we report that *CIITA* gene alterations impact survival in primary mediastinal B-cell lymphoma (PMBCL). As functional consequences of *CIITA* gene fusions, we identify downregulation of surface HLA class II expression and overexpression of ligands of the receptor molecule programmed cell death 1 (CD274/PDL1 and CD273/PDL2). These receptor–ligand interactions have been shown to impact anti-tumour immune responses in several cancers³, whereas decreased MHC class II expression has been linked to reduced tumour cell immunogenicity⁴. Thus, our findings suggest that recurrent rearrangements of *CIITA* may represent a novel genetic mechanism underlying tumour–microenvironment interactions across a spectrum of lymphoid cancers.

In Hodgkin lymphoma, translocations or chromosomal breakpoints have only rarely been described, whereas in PMBCL no recurrent translocation events have been reported⁵. Massively parallel, paired-end sequencing of expressed transcripts (RNA-seq) provides an analytical platform suitable for genome-wide mapping of translocation breakpoints, sequence variants and quantitative expression^{6,7}. Thus, we used this technology to detect novel gene fusions in the two Hodgkin lymphoma cell lines KM-H2 and L428, including 82.9 million paired-end reads for KM-H2, of which 71.1 million mapped to the reference human genome (86%, 3.6 gigabases), and 61.5 million paired-end reads for L428, of which 55.5 million (90%, 2.8 gigabases) mapped to the reference genome (Supplementary Fig. 1). Using a novel gene-fusion discovery method, we obtained 14 distinct fusion transcript predictions for KM-H2 and five for L428 (Supplementary Table 1). Figure 1 shows the fusion prediction with the highest read support involving *CIITA* and an uncharacterized gene *BX648577* (*FLJ27352*/hypothetical *LOC145788*) found in KM-H2.

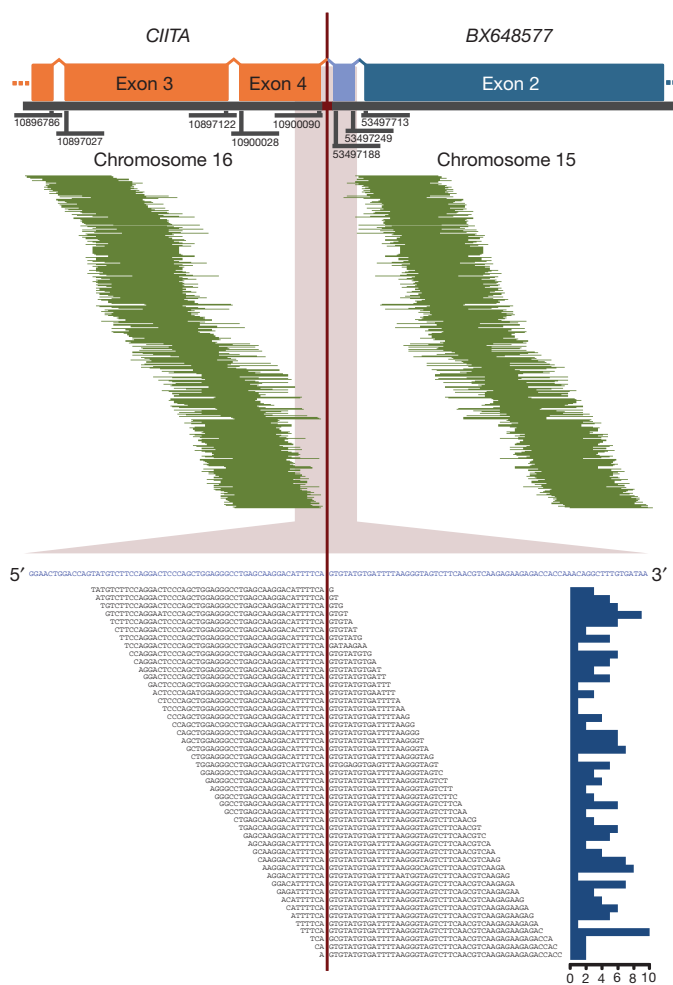


Figure 1 | *CIITA*–*BX648577* gene fusion observed using paired-end massively parallel whole transcriptome sequencing. In the upper panel, 468 mate-pair sequences are shown aligning on either side of the breakpoint (pairing *CIITA* and *BX648577*). The genomic coordinates of the exon boundaries are given. In light blue, the transcribed intronic *BX648577* sequence is shown as part of a transcript variant resulting from the fusion. In the lower panel, 191 split-reads are depicted lying on the breakpoint (in blue: merged reference sequence of *CIITA* and *BX648577*). The histogram on the right describes the absolute frequency of each sequence read spanning the breakpoint.

¹Department of Pathology and Laboratory Medicine, Centre for Lymphoid Cancers and the Centre for Translational and Applied Genomics (CTAG), Vancouver, British Columbia, V5Z4E6, Canada.

²Metabolism Branch, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland, 20892, USA. ³Department of Cell Biology and Albert Einstein Cancer Center, Albert Einstein College of Medicine, Bronx, New York, 10461, USA. ⁴Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, V5Z4S6, Canada. ⁵Department of Pathology and Medical Biology, University Medical Center Groningen, University of Groningen, 9700, The Netherlands. ⁶Division of Medical Oncology, BC Cancer Agency Centre for Lymphoid Cancer, Vancouver, British Columbia, V5Z4E6, Canada.

⁷Department of Pathology, University of Arizona, Tucson, Arizona, 85724, USA. ⁸Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, V6T1Z3, Canada.

*These authors contributed equally to this work.

Next, we validated three fusion predictions by direct sequencing and fluorescence *in situ* hybridization (FISH): the first involving the genes *BAT2L1* (chromosome 9q34.13) and *MGMT* (chromosome 10q26.3) in KM-H2 (Supplementary Fig. 2), the second involving the genes *SLCO3A1* (chromosome 15q26.1) and *ELMO1* (chromosome 7p14.1–14.2) in L428 (Supplementary Fig. 2), and lastly the previously mentioned *CIITA*–*BX648577* gene fusion (Fig. 2 and Supplementary Fig. 3). In all three cases we could confirm the breakpoint sequences as predicted. Although all of the identified gene fusions are likely to contribute to the specific phenotype of the affected cells, we focused on the *CIITA*–*BX648577* fusion transcripts given the known involvement of *CIITA* in B-cell immune function and the high read-support. The *CIITA* gene was initially studied in patients with bare lymphocyte syndrome, a rare autosomal recessive disease, in which *CIITA* mutations lead to loss of MHC class II expression and clinical manifestations due to an immunodeficiency phenotype⁸. *CIITA* was found to be the essential transactivator of MHC class II expression functioning in a complex of transcription factors (RFX, NFY, X2BP) that bind to class II MHC promoters^{9,10}.

Using PCR, the genomic breakpoint coordinates were mapped to chr15:53,489,063 and chr16:10,900,305 (NCBI Build 36.1) falling within *CIITA* exon 5 and *BX648577* intron 1 (Fig. 2a, b). Two major

transcripts of this gene fusion were identified. Sequences of the first alternative transcript aligned to exons 1–4 of *CIITA* (chromosome 16) and exon 2 of *BX648577* (chromosome 15), whereas the second alternative transcript contained an additional sequence of 62 base pairs aligning to *BX648577* intron 1, which has not been previously annotated as expressed (Fig. 2c, d). Gene expression analysis demonstrated that *BX648577* was highly overexpressed in KM-H2 compared with other Hodgkin lymphoma cell lines (35.0-fold) and microdissected germinal centre B cells (138.3-fold) (Fig. 2e, f). The longer isoform has an open reading frame consistent with a predicted 255 amino-acid protein containing the amino (N)-terminal MHC class II transactivator domains 1–4, an amino-acid sequence originating from the novel *BX648577* exon and the carboxy (C) terminus of the hypothetical *BX648577* protein (Fig. 2d and Supplementary Information).

We next performed immunofluorescence to investigate the cellular localization of the fusion protein. We found a strong perinuclear cytoplasmic staining pattern in KM-H2 compared with primary peripheral blood B cells in which we observed a predominantly nuclear localization of wild-type *CIITA* (Supplementary Fig. 4). This staining pattern is in agreement with loss of a C-terminal nuclear localization sequence¹¹ of *CIITA* in the chimaeric protein.

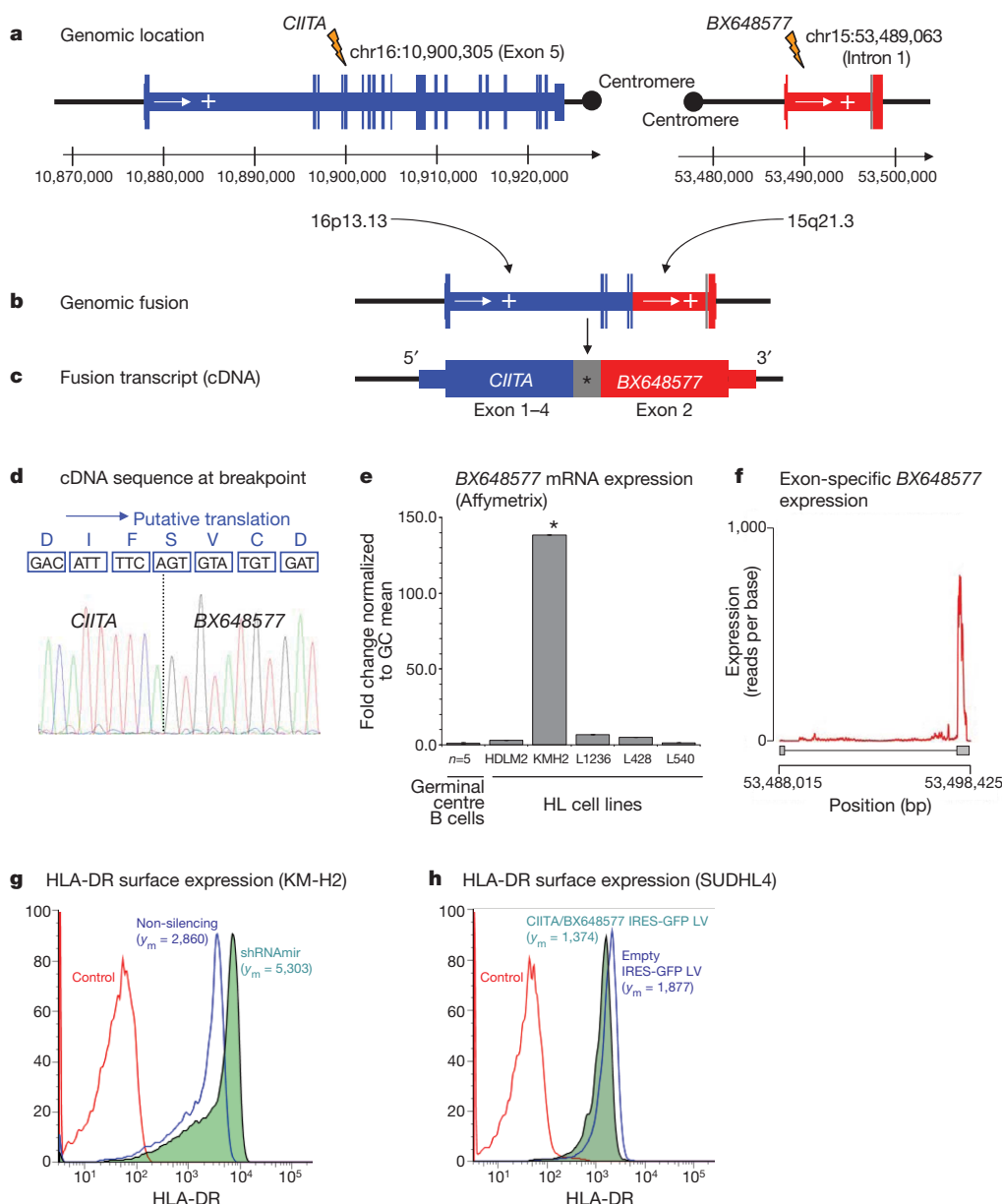


Figure 2 | Molecular characterization of the gene fusion *CIITA*–*BX648577* in Hodgkin lymphoma cell line KM-H2.

a, Genomic location: exon structure and genomic breakpoints. **b**, Genomic fusion: rearranged genomic location with fusion of *CIITA* exons 1–5 with intron 1 of *BX648577*. **c**, Fusion transcript: the longest fusion transcript with transcribed intronic *BX648577* sequence (*) is shown. Shorter splice variants exist. **d**, Reading frame at the breakpoint and putative translation: *CIITA* exons 1–4 and *BX648577* exon 2 original reading frames are conserved. The shorter splice variants leading to premature translational stop at the breakpoint are not shown. **e**, *BX648577* gene expression: gene expression array data (Affymetrix HG UA133 2.0 Plus probe set ID 243309_at) showing overexpression of *BX648577* in KM-H2 compared with microdissected germinal centre B cells and other Hodgkin lymphoma cell lines. **f**, *BX648577* exon-specific expression is biased towards exon 2 as part of the *CIITA*–*BX648577* gene fusion. **g**, RNA interference with the gene fusion in KM-H2 cells increases surface HLA-DR expression compared with the non-silencing control. **h**, Forced expression of the *CIITA*–*BX648577* fusion decreases surface HLA-DR expression on SUDHL4 cells compared with empty vector controls. Mean fluorescence intensities (y_m) are indicated.

Using high-resolution single nucleotide polymorphism analysis, we found that the rearrangement involving chromosomes 15 and 16 was associated with relative genomic imbalances at the breakpoints (Supplementary Fig. 5). Multicolour-FISH in representative metaphases of KM-H2 and interphase-FISH showed complex chromosomal rearrangements with various marker chromosomes and subclonal variation as shown by others for this cell line¹² (Supplementary Fig. 6). Locus-specific FISH using Fosmid probes for *CIITA* and *BX648577* confirmed co-localization of the *CIITA* and *BX648577* gene loci on two derivative chromosomes 16, both harbouring insertions of chromosome 15 material. These complex rearrangements are likely to place the fusion gene in a different positional context that might add to the high promoter activity of *CIITA* driving expression of the fusion transcripts in this cell line.

To investigate the functional consequence of the *CIITA*–*BX648577* gene fusion in KM-H2 cells, we generated stable KM-H2 fusion transcript knockdowns by RNA interference. Quantitative reverse transcriptase PCR (RT-PCR) showed reduction of fusion transcript levels by 85%, and western blotting demonstrated a marked reduction in *CIITA*–*BX648577* fusion protein compared with non-silencing controls (Supplementary Fig. 7). As *CIITA* is an essential transactivator of MHC class II expression⁹ and deletion mutants have been described that inhibit wild-type *CIITA* function *in vitro*^{13,14}, we studied transcriptional changes of short hairpin RNA (shRNA)-transduced cells using gene expression microarrays (Supplementary Table 2). Most strikingly, overrepresentation analysis identified genes of the ‘antigen presentation pathway’ significantly enriched among the upregulated genes ($P = 0.03$), and flow cytometry confirmed increased surface expression of HLA-DR in KM-H2 cells with fusion transcript knockdown (Fig. 2g). We also transduced diffuse large B-cell lymphoma (DLBCL) cell line SUDHL4 with a lentiviral vector containing the full-length coding complementary DNA (cDNA) of the *CIITA*–*BX648577* fusion (second alternative transcript), which led to decreased surface HLA-DR expression (Fig. 2h). These data demonstrate that the *CIITA*–*BX648577* fusion suppresses expression of HLA class II genes in

KM-H2 and SUDHL4. Diminished expression of HLA class II in B-cell lymphomas has been linked to reduced immunogenicity, escape from immunosurveillance and inferior survival^{15,16}.

To determine if breakpoints within *CIITA* are recurrent in cases of primary Hodgkin lymphoma, we used FISH on tissue microarrays and found eight out of 55 cases (15%) rearranged (Supplementary Table 3). Representative FISH images of the break-apart assay are shown in Fig. 3, demonstrating involvement of all cells of the malignant cell compartment (Hodgkin Reed–Sternberg cells that carry the CD30 antigen). In four of the eight cases we detected an unbalanced rearrangement. Because PMBCL shares clinical and biological features with cHL, we also studied PMBCL and found 29 (38%) of 77 cases positive for break-apart of the *CIITA* locus. Only four out of 131 studied cases of DLBCL (3%) showed break-apart of the *CIITA* locus. Hence, the occurrence of *CIITA* rearrangements in PMBCL and cHL was significantly higher than in DLBCL ($P < 0.0001$).

We also studied the correlation of *CIITA* translocation with clinical outcome in 57 patients with PMBCL who were treated with curative intent using multi-agent chemotherapy with or without radiotherapy¹⁷. The clinical characteristics of this cohort are summarized in Supplementary Table 4. Presence of a *CIITA* rearrangement significantly correlated with a shorter disease-specific survival (10-year disease-specific survival 63.6% compared with 85.0%, $P = 0.044$) (Fig. 3d). Multivariate testing using a Cox regression model including the International Prognostic Index¹⁸ score and the individual clinical prognostic factors showed that the presence of *CIITA* rearrangement had independent prognostic significance ($P = 0.013$) for decreased disease-specific survival together with elevated LDH serum levels ($P = 0.001$) (Supplementary Table 5).

To determine fusion partners involved in *CIITA* rearrangements, we used 3′ rapid amplification of cDNA ends (RACE) in eight cases of PMBCL that were positive for *CIITA* translocation by FISH. In total, we identified five new translocation partners (*CD274*, *CD273*, *RALGDS*, *RUNDC2A* and *C16ORF75*) (Supplementary Table 6). Precise mapping revealed that the corresponding genomic breakpoints in *CIITA* all

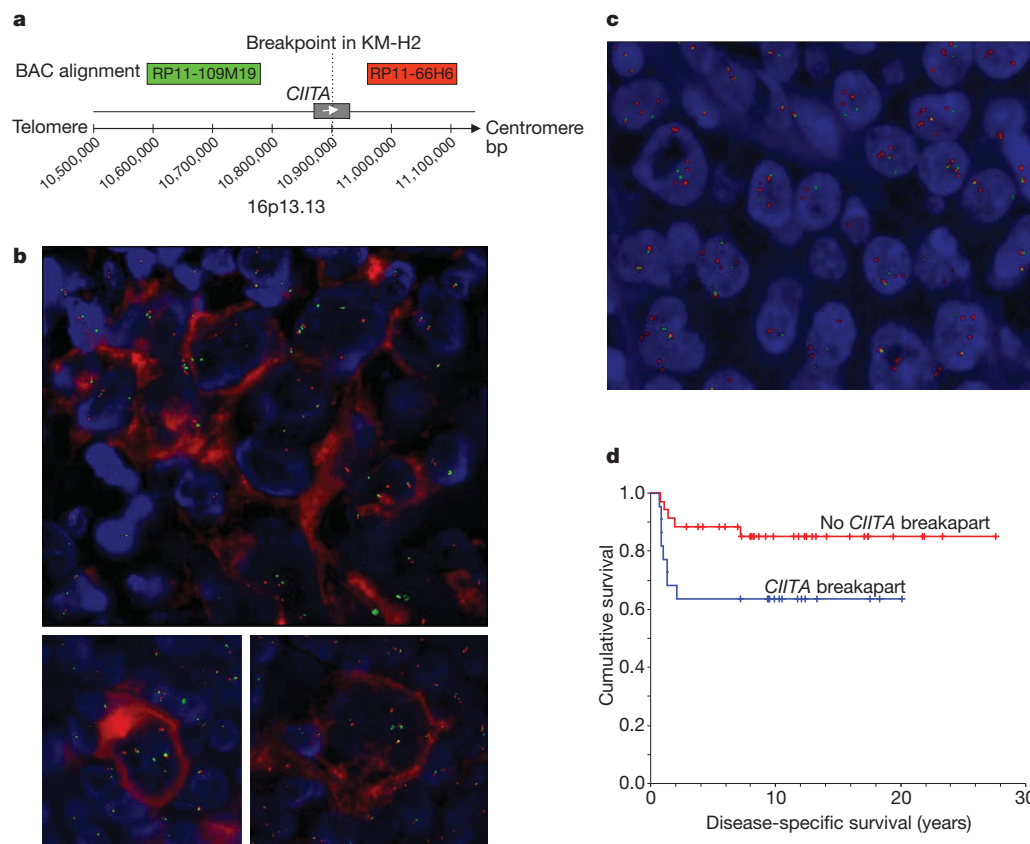


Figure 3 | FISH on tissue microarrays showing recurrent *CIITA* break-apart in cHL and PMBCL. Representative images are shown. **a**, Design of the break-apart assay using bacterial artificial chromosome probes RP11-109M19-SpG (green signals) plus RP11-66H6-SpO (red signals). **b**, Combined immunofluorescence for CD30 (red staining) and FISH shows *CIITA* rearrangement in Hodgkin Reed–Sternberg cells that carry the CD30 antigen. Upper image: Hodgkin Reed–Sternberg cell-rich area with *CIITA* break-apart, lower images: individual Hodgkin Reed–Sternberg cells with break-apart. **c**, PMBCL case with *CIITA* break-apart in almost all cells represented in the section. Signal constellation indicates *CIITA* polyploidy and rearrangement of multiple alleles. **d**, Disease-specific survival of 57 patients with PMBCL treated with multi-agent chemotherapy (with or without radiation) according to *CIITA* rearrangement status. The presence of a *CIITA* rearrangement significantly correlated with shorter disease-specific survival ($P = 0.044$).

fell within a 1.6-kilobase (kb)-spanning breakpoint cluster region in intron 1, indicating a 'hot spot' area for chromosomal breaks (Supplementary Fig. 8). In the case of *CD274* (*PDL1*), *CD273* (*PDL2*) and *RUNDC2A*, the fusion transcripts merged exon 1 of *CIITA* with exon 2 of the respective fusion partners, resulting in open reading frames (Supplementary Information). Of note, in both cases with *CIITA*-*CD273* fusions, mutations to the *CIITA* start codon were detected resulting in translational reading frames beginning at the original *CD273* start site (Supplementary Information). Overexpression of *CD274* and *CD273* closely linked to copy number gain has been previously described in PMBCL and Hodgkin lymphoma^{19–21}, and

expression of *CD274* has been shown to correlate with poor prognosis in other cancers²². Furthermore, re-analysis of gene expression profiling data confirmed overexpression of both genes in over 50% of cases of PMBCL (Fig. 4a)²³. Here, we could show by quantitative RT-PCR that in all three cases *CD274* and *CD273* were highly overexpressed (fold changes 451, 4,115 and 1,729, respectively) as part of the respective *CIITA* fusions compared with *CD77*⁺ centroblasts (Fig. 4b). Importantly, expression levels markedly exceeded the levels of those in cases of PMBCL without translocations. Overall, these results indicate that substitution of active *CIITA* promoters²⁴ can lead to overexpression of partner genes in a B-cell context.

Next, we sought to study the functional consequences of *CD274* and *CD273* wild-type overexpression. We selected cell line U2940 (ref. 25) (consistent with PMBCL features) expressing high levels of both *CD274* and *CD273* messenger RNA (mRNA) and membrane protein (Fig. 4b, c). To study the effects of *CD274* and *CD273* surface expression on T-cell activation in particular, we analysed a co-culture system in which Jurkat T cells were co-incubated with U2940 cells. First, we confirmed that PD-1 expression was induced in Jurkat T cells upon T-cell receptor stimulation (Supplementary Fig. 9a). After 4 h of co-culture, Jurkat T-cell activation, measured by CD69 expression, was markedly decreased with increasing admixture of U2940 cells (Fig. 4d). In contrast, inhibition of T-cell activation was not observed when U2932 and SUDHL4 cells were co-incubated that did not express *CD274* or *CD273* on their surface (Supplementary Fig. 9b). We could also show that inhibition of T-cell activation by U2940 cells was reversible by incubation with inhibitors against PD-1, *CD274* and *CD273*, but not against CTLA4 (Fig. 4e). Furthermore, overexpression of PD-1 in Jurkat cells further augmented T-cell inhibition, indicating dependence on the PD-1 pathway (Supplementary Fig. 9c).

We also investigated the effects of forced expression of the *CIITA*-*CD274* and *CIITA*-*CD273* gene fusions in U2932 cells, which do not express wild-type PD-1 ligands on their surface. Ectopic expression of the full-length coding cDNA of the fusions in U2932 markedly increased surface expression of the respective proteins in transduced cells (Fig. 4f) and, similar to U2940, decreased CD69 and interleukin-2 (IL-2) expression of PD-1-overexpressing Jurkat T cells compared with empty vector controls. In summary, these data show that both wild-type PD-1 ligand expression and the identified *CIITA*-*CD274*/*CD273* gene fusions negatively regulate Jurkat T-cell activation.

Taken together, we have discovered breakpoints in the master regulator of MHC class II expression *CIITA* that are novel and highly

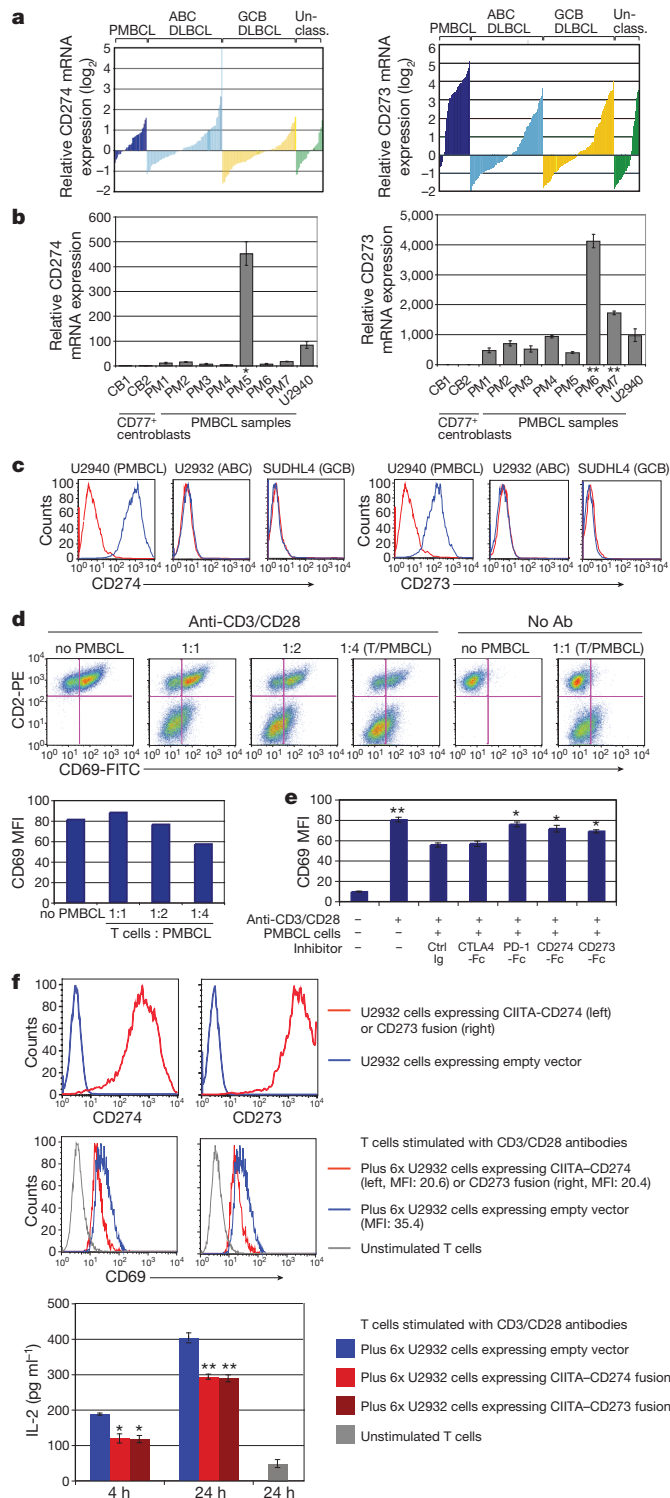


Figure 4 | CD274/CD273 expression on PMBCL cells inhibits T-cell activation. **a**, mRNA overexpression of *CD274* and *CD273* in molecular subtypes of DLBCL including PMBCL. Normalized relative log₂ ratios are shown (Affymetrix gene expression profiling)²³. Probeset intensities of *CD274* and *CD273* correlate with each other (Pearson coefficient 0.541). **b**, Fold change of *CD274* and *CD273* mRNA expression in PMBCL compared with germinal centre B cells including cases with *CIITA*-*CD274* (*), *CIITA*-*CD273* (**) fusions and cell line U2940. **c**, Flow cytometric analysis of *CD274*/*CD273* expression. Expression levels of *CD274* or *CD273* (blue histogram) and isotype controls (red histogram) are shown. **d**, Inhibition of T-cell activation by PMBCL cells. Dot plots and bar graph show that CD69 expression on Jurkat T cells (CD2 positive) is increasingly reduced by co-incubation with increasing numbers of U2940 cells. **e**, T-cell inhibition is PD1 dependent. *CD274*-Fc, *CD273*-Fc, PD-1-Fc, CTLA4-Fc or control Ig were added to the co-cultures, respectively, in which Jurkat T cells were mixed with U2940 cells. Mean fluorescence intensities (MFI) are shown. Each bar is the mean of triplicate cultures, the error bars indicating the standard deviation (Student's *t*-test, **P* < 0.05, ***P* < 0.01, compared with control Ig). **f**, Forced expression of *CIITA*-*CD274* and *CIITA*-*CD273* fusion transcripts in U2932. The top panel shows the surface protein expression in transduced cells compared with empty vector control cells. The middle and bottom panels show significantly reduced CD69 expression on Jurkat T cells and IL-2 levels in the supernatant after admixture of *CIITA*-*CD274* and -*CD273* fusion-expressing U2932 cells. Each bar is the mean of triplicate cultures, the error bars indicating the standard deviation (Student's *t*-test, **P* < 0.05, ***P* < 0.01, compared with the empty vector control).

recurrent in PMBCL. Moreover, the occurrence of *CIITA* breaks in cHL further substantiates the relatedness of cHL and PMBCL^{19,26}. We observed *CIITA* breaks in only 3% of cases of DLBCL, contrasting with 38% in PMBCL, which highlights the relative specificity of this genetic event. In PMBCL, *CIITA* breaks were significantly correlated with outcome, demonstrating the potential clinical importance of these rearrangements, a finding that will require validation in additional cohorts of patients. Consequences of *CIITA* rearrangements appear to be diverse as evidenced by multiple fusion partners, and concomitant chromosomal imbalances as described previously in PMBCL. However, functional study of the most frequent specific gene fusions suggests that escape from immunosurveillance through various mechanisms might play an important role in the pathogenesis of these lymphomas. Furthermore, our data raise the possibility that deletion of tumour suppressor genes²⁷, overexpression of oncogenes resulting from gene fusion, and *CIITA* loss of function might be concurrent consequences of a single genetic event.

METHODS SUMMARY

RNA-seq data were generated from the Hodgkin lymphoma cell lines KM-H2 and L428 as previously described²⁸, and fusion transcripts were detected using deFuse (<http://compbio.bccrc.ca>). The validation of fusion transcripts was performed by FISH and PCR. Protein expression of the *CIITA*-BX648577 fusion protein was detected by western blotting and immunofluorescence. Fusion transcript knock-down in KM-H2 cells was performed by lentiviral transduction of a vector expressing a shRNA interfering with BX648577 sequences. Forced expression of the *CIITA* gene fusions was achieved using a lentiviral²⁹ and an inducible retroviral vector³⁰. Gene expression in KM-H2 cells was measured using gene expression profiling and flow cytometry. Primary paraffin tissue samples of 263 patients with B-cell lymphoma were studied using FISH. For eight selected cases of PMBCL with *CIITA* rearrangements, we performed 3' RACE. Quantitative RT-PCR was performed to measure expression of *CIITA*-BX648577 fusion transcripts, *CD274* and *CD273*. For *in vitro* studies we used U2940, U2932 and SUDHL4 cells. We used Jurkat T cells for co-culture experiments measuring T-cell activation by CD69 flow cytometric analysis and IL-2 enzyme-linked immunosorbent assays (ELISA).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 April; accepted 14 December 2010.

Published online 2 March 2011.

- Dyer, M. J. The pathogenetic role of oncogenes deregulated by chromosomal translocation in B-cell malignancies. *Int. J. Hematol.* **77**, 315–320 (2003).
- Kerckaert, J. P. *et al.* LAZ3, a novel zinc-finger encoding gene, is disrupted by recurring chromosome 3q27 translocations in human lymphomas. *Nature Genet.* **5**, 66–70 (1993).
- Blank, C., Gajewski, T. F. & Mackensen, A. Interaction of PD-L1 on tumor cells with PD-1 on tumor-specific T cells as a mechanism of immune evasion: implications for tumor immunotherapy. *Cancer Immunol. Immunother.* **54**, 307–314 (2005).
- Rimsza, L. M. *et al.* Loss of MHC class II gene and protein expression in diffuse large B-cell lymphoma is related to decreased tumor immunosurveillance and poor patient survival regardless of other prognostic factors: a follow-up study from the Leukemia and Lymphoma Molecular Profiling Project. *Blood* **103**, 4251–4258 (2004).
- Martin-Subero, J. I. *et al.* Chromosomal breakpoints affecting immunoglobulin loci are recurrent in Hodgkin and Reed-Sternberg cells of classical Hodgkin lymphoma. *Cancer Res.* **66**, 10332–10338 (2006).
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621–628 (2008).
- Palanisamy, N. *et al.* Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma. *Nature Med.* **16**, 793–798 (2010).
- Reith, W. & Mach, B. The bare lymphocyte syndrome and the regulation of MHC expression. *Annu. Rev. Immunol.* **19**, 331–373 (2001).
- Silacci, P., Mottet, A., Steimle, V., Reith, W. & Mach, B. Developmental extinction of major histocompatibility complex class II gene expression in plasmacytes is mediated by silencing of the transactivator gene *CIITA*. *J. Exp. Med.* **180**, 1329–1336 (1994).
- Masternak, K. *et al.* *CIITA* is a transcriptional coactivator that is recruited to MHC class II promoters by multiple synergistic interactions with an enhanceosome complex. *Genes Dev.* **14**, 1156–1166 (2000).
- Cressman, D. E., Chin, K. C., Taxman, D. J. & Ting, J. P. A defect in the nuclear translocation of *CIITA* causes a form of type II bare lymphocyte syndrome. *Immunity* **10**, 163–171 (1999).
- Joos, S. *et al.* Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including *REL* and *JAK2*. *Int. J. Cancer* **103**, 489–495 (2003).
- Chin, K. C., Li, G. & Ting, J. P. Activation and transdominant suppression of MHC class II and HLA-DMB promoters by a series of C-terminal class II transactivator deletion mutants. *J. Immunol.* **159**, 2789–2794 (1997).
- Zhou, H., Su, H. S., Zhang, X., Douhan, J., III & Glimcher, L. H. *CIITA*-dependent and -independent class II MHC expression revealed by a dominant negative mutant. *J. Immunol.* **158**, 4741–4749 (1997).
- Roberts, R. A. *et al.* Loss of major histocompatibility class II gene and protein expression in primary mediastinal large B-cell lymphoma is highly coordinated and related to poor patient survival. *Blood* **108**, 311–318 (2006).
- Diepstra, A. *et al.* HLA class II expression by Hodgkin Reed-Sternberg cells is an independent prognostic factor in classical Hodgkin's lymphoma. *J. Clin. Oncol.* **25**, 3101–3108 (2007).
- Savage, K. J. *et al.* Favorable outcome of primary mediastinal large B-cell lymphoma in a single institution: the British Columbia experience. *Ann. Oncol.* **17**, 123–130 (2006).
- The International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N. Engl. J. Med.* **329**, 987–994 (1993).
- Rosenwald, A. *et al.* Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. *J. Exp. Med.* **198**, 851–862 (2003).
- Yamamoto, R. *et al.* PD-1-PD-1 ligand interaction contributes to immunosuppressive microenvironment of Hodgkin lymphoma. *Blood* **111**, 3220–3224 (2008).
- Green, M. R. *et al.* Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood* **116**, 3268–3277 (2010).
- Thompson, R. H. *et al.* Costimulatory B7-H1 in renal cell carcinoma patients: Indicator of tumor aggressiveness and potential therapeutic target. *Proc. Natl Acad. Sci. USA* **101**, 17174–17179 (2004).
- Lenz, G. *et al.* Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl Acad. Sci. USA* **105**, 13520–13525 (2008).
- Lennon, A. M. *et al.* Isolation of a B-cell-specific promoter for the human class II transactivator. *Immunogenetics* **45**, 266–273 (1997).
- Sambade, C. *et al.* U-2940, a human B-cell line derived from a diffuse large cell lymphoma sequential to Hodgkin lymphoma. *Int. J. Cancer* **118**, 555–563 (2006).
- Savage, K. J. *et al.* The molecular signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-cell lymphomas and shares features with classical Hodgkin lymphoma. *Blood* **102**, 3871–3879 (2003).
- Melzner, I. *et al.* Biallelic deletion within 16p13.13 including SOCS-1 in Karpas 1106P mediastinal B-cell lymphoma line is associated with delayed degradation of JAK2 protein. *Int. J. Cancer* **118**, 1941–1944 (2006).
- Morin, R. *et al.* Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques* **45**, 81–94 (2008).
- Steidl, U. *et al.* A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. *J. Clin. Invest.* **117**, 2611–2620 (2007).
- Ngo, V. N. *et al.* A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* **441**, 106–110 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work is supported by a postdoctoral fellowship of the Cancer Research Society (Steven E. Drabin Fellowship) to C.S., the Michael Smith Foundation for Health Research to C.S. and S.P.S., the Lymphoma Research Foundation to C.S. and the Canadian Breast Cancer Foundation to S.P.S. Operational funds were available through the Canadian Institutes of Health Research, grant number 178536 to R.D.G. R.D.G., J.M.C., M.A.M. and D.E.H. are also supported by the Terry Fox Foundation (number 019001). This work was in part supported by an infrastructure grant of Genome Canada/Genome BC. U.S. is the recipient of a Howard Temin Award of the National Institutes of Health/National Cancer Institute (R00CA131503), a new investigator award of the Leukemia Research Foundation, and is the Diane and Arthur B. Belfer Faculty Scholar in Cancer Research of the Albert Einstein College of Medicine. We thank G. Simkin, C. Polumbo and T. Vogler for technical support. L.R. is the recipient of a CJ Martin Fellowship from the National Health and Medical Research Council of Australia.

Author Contributions C.S. designed the research, performed FISH, PCR and direct sequencing, interpreted results and wrote the paper. S.P.S. designed the research, analysed the transcriptome data and wrote the paper. B.W.W. performed PCR and interpreted results. M.K., U.C.O. and L.R. performed *in vitro* functional analyses. P.F. reviewed pathology and constructed the tissue microarrays. N.A.J. performed single nucleotide polymorphism analyses. Y.Z. performed library construction and RNA-seq. A.T. performed nucleic acid extraction and quantitative RT-PCR. B.M. and S.B.N. performed FISH. A.M., M.S., G.L. and S.J.J. analysed the transcriptome data. A.D., A.B., L.R. and D.E.H. interpreted results. J.M.C. and K.J.S. provided clinical data. D.G.H. designed the research. L.M.S. and U.S. designed the research and interpreted results. M.A.A. designed the research. R.D.G. designed the research, constructed the tissue microarrays, interpreted results and wrote the paper.

Author Information Array data are deposited in NCBI Gene Expression Omnibus under accession number GSE25990. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.D.G. (rgascoyn@bccancer.bc.ca).

METHODS

Materials and patient samples. The Hodgkin lymphoma cell lines KM-H2 and L428 were selected for whole transcriptome paired-end sequencing (RNA-seq) analysis. Both cell lines originated from patients with cHL (Epstein Barr virus negative mixed cellularity and nodular sclerosis, respectively). Additionally, cHL cell lines L540, L1236, HDLM2 and microdissected germinal centre cells were used for comparison of gene expression levels. All cell lines were obtained from the German Collection of Microorganisms and Cell Cultures and cultures were grown following standard recommendations (<http://www.dsmz.de>). DNA and RNA were extracted using Allprep extraction kits (Qiagen).

For validation in primary tissue samples, we studied paraffin material from 263 patients using FISH. All cases were identified through the Lymphoid Cancer Database of the British Columbia Cancer Agency including 131 patients with DLBCL, 77 with PMBCL and 55 with cHL. Patients with PMBCL were treated with CHOP and CHOP-like chemotherapy as previously described¹⁷. Additionally, we studied eight cases of PMBCL using RACE. All cases included in this study were independently reviewed by two haematopathologists (R.D.G. and P.F.) before analysis. Ethical approval for this study was obtained from the University of British Columbia – British Columbia Cancer Agency Research Ethics Board. Previously published gene expression profiling data of 203 biopsy samples with defined molecular subtypes of DLBCL, including PMBCL, were re-analysed²³.

For functional *in vitro* studies, we used cell line U2940, derived from a DLBCL with features consistent with PMBCL²⁵, U2932 (DLBCL, activated B-cell phenotype)³¹ and SUDHL4 (DLBCL, germinal centre B-cell phenotype)³². Jurkat T cells were used for co-culture experiments measuring T-cell activation.

Whole transcriptome paired-end sequencing (RNA-seq). RNA-seq was performed as previously described^{28,33}. In brief, double-stranded cDNA was synthesized from polyadenylated RNA and sheared. The 190–210 base-pair fraction was isolated and amplified with ten cycles of PCR using the Illumina Genome Analyser paired-end library protocol (Illumina). The resulting libraries were then sequenced on an Illumina Genome Analyser II.

Short read sequences obtained from the Genome Analyser II were mapped to the reference human genome (NCBI build 36.1, hg18, <http://genome.ucsc.edu/cgi-bin/hgGateway>), and spliced (cDNA) and unspliced gene sequences (Ensembl version 54) were aligned using the Bowtie algorithm³⁴. Fusion transcript discovery was performed using deFuse (<http://compbio.bccrc.ca>). Briefly, deFuse identifies fusion transcripts by clustering discordantly aligning paired-end reads as potential evidence of reads that span a fusion breakpoint. Clusters of discordantly aligning reads are then used to inform a targeted search for reads split by fusion breakpoints. The results produced by deFuse were further filtered to reduce the number of false positives: (1) predictions had to be supported by at least eight reads spanning a fusion breakpoint and five reads split by a fusion breakpoint; (2) fusions between adjacent genes were removed unless implied in genomic inversion or eversion; (3) predictions involving ribosomal proteins or small nuclear ribosomal proteins were removed; (4) the results were filtered against fusion predictions from 41 ovarian cancer libraries in an attempt to remove systematic technical artefacts. After these series of filters were applied, a candidate list of fusions was carried forward for further analysis.

Gene expression levels were determined according to the total coverage of a gene, which was defined as the sum of the coverage of each non-redundant exonic nucleotide normalized by total mapped nucleotides.

Gene expression profiling of Hodgkin lymphoma cell lines. RNA-seq data were compared with gene expression data of Hodgkin lymphoma cell lines (KM-H2, L428, L540, L1236 and HDLM2) and microdissected germinal centre cells ($n = 5$) derived from Affymetrix HG 2.0 Plus gene expression arrays using two-cycle labelling reactions according to the standard protocol. For comparison of KM-H2 knockdown cultures (see below) we used freely available software (DCHIP, <http://biosun1.harvard.edu/complab/dchip>). Overrepresentation and pathway analysis was performed using Ingenuity Pathway Analysis (Ingenuity Systems).

High-resolution single nucleotide polymorphism analysis. Genome-wide copy number analysis was performed using Human SNP 6.0 arrays (Affymetrix) following the standard protocol and the manufacturer's instructions. Hybridized arrays were washed, stained and scanned using Affymetrix Fluidics Station 450 and an Affymetrix GeneChip Scanner. The Affymetrix SNP 6.0 data were processed using the R-based *aroma.affymetrix* package.³⁵ Total copy number was estimated according to the CRMA procedure, calibrating for allelic cross talk, nucleotide-position probe sequence effects, robust probe summarization and correction for fragment length effects. CRMA-normalized data were then compared with a reference derived from the HapMap 270 set of arrays downloaded from the Affymetrix website. Breakpoints were determined by a modified version of a hidden Markov model described before³⁶. Software (including visualization routines) is available as part of the CNA-HMMer package available at <http://compbio.bccrc.ca>.

Genomic and RT-PCR and FISH. The validation of fusion transcripts was performed using both genomic and RT-PCR with forward and reverse primer combinations designed within the margins of the paired-end read sequences detected by RNA-seq (Supplementary Table 7). Expressed fusion transcript variants were cloned into pCR 2.1-TOPO vectors for amplification in chemically competent *Escherichia coli* (TOPO TA cloning kit, Invitrogen). Consensus sequences were determined from at least five separate colonies. For further validation by FISH, fusion probes were used to confirm gene fusions in interphase and metaphase nuclei of KM-H2 and L428. For probe design see Supplementary Fig. 10a–c. For detailed cytogenetic characterization of KM-H2, we used multicolour-FISH using 24Xcyte multi-colour probes (MetaSystems) as previously described³⁷ and locus-specific FISH using Fosmid probes available from the Children's Hospital & Research Center of Oakland for the gene loci of *CIITA* (WI2-1388110, G248P83694E5) and *BX648577* (WI2-2329H16, G248P88223D8).

Lentiviral transductions. RNA interference with the *CIITA*-*BX648577* fusion transcript was performed using lentiviral transduction of a vector (pGIPZ-sh.FLJ27352, clone V2LHS_212659, Thermo Scientific Open Biosystems) expressing an shRNA (TGCTGTTGACAGTGAGCGCGACAGCCACCTCACTATCAAAATAGTGAAGCCACAGATGTATTTGATAGTGAGGTGGCTGCTTGGCTACTGCCTCGGA) that, after processing to mature siRNA, interferes with *BX648577* exon 2 sequence as part of the *CIITA*-*BX648577* fusion transcript. A vector with a non-interfering shRNA insert (non-silencing control) was used for comparison. In brief, the shRNAmir constructs were co-transfected along with the helper and packaging plasmids (pHDM-G, pHDM-Hgpm2, pHDM-tat1b and pRC/CMV-rev1b) into 293T cells by calcium-phosphate-mediated transfection (CalPhos Mammalian Transfection kit (Clontech)). Supernatant was collected 48 h and 72 h after transfection and concentrated by ultracentrifugation for 2 h at 60,000g, at 4 °C. The lentiviral particles were resuspended in IMDM and then used for transduction of KM-H2 cells. KM-H2 parental cells were transduced by adding 8 $\mu\text{g ml}^{-1}$ polybrene, with both pGIPZ-sh.FLJ27352 and non-silencing control virus, and green fluorescent protein (GFP)-positive cells were sorted on a BD FACSAria cell sorter (BD Biosciences) 3 days after transduction and cultured using standard recommendations (<http://www.dsmz.de>). Purity (GFP positivity) was greater 93% for all experiments. Efficiency of RNA interference was evaluated by measuring residual expression of the fusion transcript by quantitative RT-PCR (see Supplementary Table 7 for primer sequences). Samples were run in triplicate on an iQ5 real-time PCR instrument (BioRad) under standard conditions using SYBR Green (Applied Biosystems) for detection. Measurements were quantified using the $\Delta\Delta\text{CT}$ method (Pfaffl) and expressed relative to the expression in parental (wild type) KM-H2 cells.

For forced expression of the *CIITA*-*BX648577* gene fusion, we PCR-amplified the full-length cDNA from KM-H2 cells, and cloned it into the pGEM-T easy TA cloning vector (Promega). The fusion gene was then transferred into the pCAD-IRES-GFP lentiviral vector²⁹. An empty vector was used as a negative control. For lentivirus production, we co-transfected lentiviral vectors and helper plasmids into 293T cells, harvested the supernatant and concentrated the lentiviral particles by ultracentrifugation (L7-65, Beckman) at 60,000g, 2 h, 4 °C. The resultant concentrated lentiviral suspension was then used to transduce SUDHL4 target cells. Three to five days after transduction, we sorted GFP-positive cells, which were used for subsequent experiments.

Western blotting and immunofluorescence. Semi-quantitative measurement of fusion protein expression was performed by semi-dry western blotting using a mouse monoclonal anti-*CIITA* primary antibody (7-1H, Santa Cruz Biotechnology) directed against the N terminus of *CIITA*, and a goat anti-mouse IgG-HRP secondary antibody (Santa Cruz). An anti- β -actin goat polyclonal antibody (C-11) followed by a donkey anti-goat IgG-HRP (both Santa Cruz) served as a control. RNA extraction of cell cultures and gene expression profiling using Affymetrix GeneChip HG 133 2.0 plus arrays was performed as previously described³⁸.

KM-H2 cells and primary human B cells (CD20⁺) were attached to poly-lysine-coated slides. To fix the cells, we incubated them with 4% PFA for 10 min at room temperature and treated with 0.15% Triton X in PBS for 2 min at room temperature. After blocking, cells were incubated with an anti-*CIITA* antibody (clone 7-1H, Santa Cruz Biotechnology) overnight at 4 °C. After washing, cells were incubated with Alexa Fluor 594 rabbit anti-mouse IgG (Invitrogen) for 45 min at 37 °C. After stringent washing, cells were counterstained with DAPI and covered. Slides were analysed on a laser confocal microscope (Leica SP5 AOBs).

FISH on tissue microarrays. Tissue microarrays were constructed using archival, formalin-fixed, paraffin-embedded diagnostic biopsy specimens and 0.6-mm duplicate cores for DLBCL and PMBCL cases, and 1.5-mm cores for cHL, respectively. Optimal areas of the biopsies were chosen containing frequent large B or Hodgkin Reed–Sternberg cells, respectively. FISH was performed as previously described³⁹ using in-house bacterial artificial chromosome break-apart and fusion probes for the loci of *CIITA* and *BX648577* (Supplementary Fig. 10a).

For fluorescence immunophenotyping and interphase cytogenetics as a tool for the investigation of neoplasms (FICTION) experiments, we used a primary monoclonal mouse CD30 antibodies (Ber-H2, Dako) and Alexa Fluor 594, goat anti-mouse (IgG, Invitrogen, Molecular Probes), following a standard protocol as previously described⁴⁰. Commercially available Vysis LSI BCL6 Dual Colour, Break Apart Rearrangement probes (Abbott Molecular) were used in selected samples. All cases were independently scored by C.S. and S.B.N. For DLBCL and PMBCL, cases were recorded as rearranged for a certain locus if break-apart occurred in more than 5% of nuclei in three different high-power fields. All other signal constellations were regarded as negative. cHL cases were scored positive if at least ten cells with a signal split were detected in typically large cells distributed across the biopsy core. Unbalanced rearrangements were recorded if unequal numbers of red or green signals were seen in most nuclei.

RACE and quantitative RT-PCR. We used 3' RACE to identify aberrant transcript variants of *CIITA* in selected PMBCL samples. Total RNA was extracted from eight cases with PMBCL with fresh-frozen lymph node material available after mechanical homogenization using Allprep extraction kits (Qiagen). cDNA was prepared from 150 ng to 1 µg of total RNA using the SMARTer RACE cDNA Amplification Kit (Clontech Laboratories) following the manufacturer's protocol. 3' RACE PCR amplification of the cDNA was performed using gene-specific primers for *CIITA* exon 1 and Universal Primer Mix. Aberrant length products were gel-purified and sequenced directly or were cloned into pCR 2.1-TOPO vectors (TOPO TA cloning kit, Invitrogen) and fully sequenced. Chimeric mRNAs were confirmed by RT-PCR of randomly primed cDNA of the respective cases using gene-specific primers.

Quantitative RT-PCR was performed to measure expression levels of *CD274* and *CD273* mRNA using an Applied Biosystems 7900HT real-time PCR system. Primers were designed for SYBR Green (Applied Biosystems) detection of both wild-type and fusion transcripts for each gene (Supplementary Table 7). RNA was extracted from cell line U2940 and seven PMBCL lymph node samples (cell disaggregates) using Allprep extraction kits (Qiagen). PMBCL cases included one with previously identified *CIITA*-*CD274* fusion and two with *CIITA*-*CD273* fusions. As expression controls, magnetically enriched CD77⁺ cells were used (benign tonsillar cell disaggregates) (Miltenyi Biotec) as previously described⁴¹.

Flow cytometric analysis and ELISA. U2940, U2932 and SUDHL4 cell pellets were washed in 1 × PBS with 1% FCS buffer and incubated with antibodies against CD69, CD2, CD274, CD273 and PD-1 (all from BD Pharmingen). Jurkat T cells were stimulated with anti-CD3 and anti-CD28 antibodies (BD Pharmingen). Recombinant human PD-1-Fc, CD274-Fc, CD273-Fc, CTLA4-Fc and control IgG (R&D Systems) were used for specific inhibition of surface molecules. KM-H2 and SUDHL4 cells were incubated with PE-Cy5-conjugated anti-human HLA-DR antibody (Beckman Coulter), and analysed on a BD FACSAria II instrument (BD Biosciences). IL-2 levels were measured using the Human IL-2 ELISA Kit (R&D Systems, number D2050) according to the manufacturer's instructions. Plates were read at 405 nm using an ELISA reader (Molecular Devices).

Retroviral transduction. PD-1 cDNA (Invitrogen) was subcloned into an pBMN-IRES-EGFP vector, which was modified from the original Moloney LTR vector pBMN-IRES-Lyt2 (provided by the G. Nolan laboratory, Stanford University) by replacing Lyt2 with EGFP.

Lipofectamine 2000 (Invitrogen) was used to transfect 293T producer cells with a plasmid mixture for gag and pol, a mutant ecotropic env and each particular retrovirus. After 2 days, supernatant was passed through a 0.45-µm filter, mixed with polybrene (8 µg ml⁻¹) and used for centrifugal transduction of target cells expressing the ecotropic retroviral receptor. In some instances a second infection was performed, using fresh supernatant collected 3 days after transfection of producer cells.

U2932 cells were first transduced with a feline endogenous virus expressing the ecotropic retroviral receptor. Ecotropic receptor-expressing cells were then transduced with a retrovirus expressing the bacterial tetracycline repressor using blasticidin as the selectable marker. Single-cell clones were screened for tetracycline repressor expression. An inducible retroviral vector was used to drive expression of *CIITA*-*CD274* or *CIITA*-*CD273* fusions (for full-length sequences see Supplementary Information) under the control of a cytomegalovirus promoter containing tetracycline repressor binding sites as described previously³⁰. A mixture of *CIITA*-*CD274*- or *CIITA*-*CD273*-containing plasmid DNA, the mutant ecotropic envelope-expressing plasmid pHIT/EA6x3* and gag-pol expressing plasmid pHIT60 was used to transfect 293T cells using the Lipofectamine 2000 reagent (Invitrogen). Two days after transfection, retrovirus supernatants were collected to infect the engineered U2932 cells in the presence of 8 mg ml⁻¹ polybrene (Sigma) in a single spin infection, and puromycin was used to select for stable integrants over 6 days. Fusion gene expression was induced by doxycycline (20 ng ml⁻¹) for 2 days before flow cytometry analysis and Jurkat T-cell co-culture experiments. For flow cytometric analysis of intracellular cleaved poly(ADP-ribose) polymerase (PARP) and active caspase 3 in U2932, cells were fixed in PBS containing 2% paraformaldehyde, then permeabilized and co-stained with anti-PARP (BD Pharmingen, clone F21-852) and anti-caspase-3 (BD Pharmingen number 51-68655X) antibodies diluted in FACS buffer containing 0.25% saponin (Sigma).

Statistical analysis. Group comparisons were performed using Student's *t*-tests and χ^2 analysis. For time-to-event analyses, we used disease-specific survival as the primary endpoint. Disease-specific survival was defined as the time from initial diagnosis to lymphoma- or treatment-related death. Time to event analyses using the Kaplan-Meier method was performed using SPSS version 11.0.0. Multivariate analysis was performed using a Cox regression model (forward stepwise) with the input variables of age, tumour stage, serum lactate dehydrogenase concentration, World Health Organization performance status, number of extranodal sites and presence of *CIITA* rearrangement.

- Amini, R. M. *et al.* A novel B-cell line (U-2932) established from a patient with diffuse large B-cell lymphoma following Hodgkin lymphoma. *Leuk. Lymphoma* **43**, 2179–2189 (2002).
- Epstein, A. L. *et al.* Biology of the human malignant lymphomas. IV. Functional characterization of ten diffuse histiocytic lymphoma cell lines. *Cancer* **42**, 2379–2391 (1978).
- Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Bengtsson, H., Wirapati, P. & Speed, T. P. A single-array preprocessing method for estimating full-resolution raw copy numbers from all Affymetrix genotyping arrays including GenomeWideSNP 5 & 6. *Bioinformatics* **25**, 2149–2156 (2009).
- Shah, S. P. *et al.* Integrating copy number polymorphisms into array CGH analysis using a robust HMM. *Bioinformatics* **22**, e431–e439 (2006).
- Lestou, V. S. *et al.* Multicolour fluorescence in situ hybridization analysis of t(14;18)-positive follicular lymphoma and correlation with gene expression data and clinical outcome. *Br. J. Haematol.* **122**, 745–759 (2003).
- Steidl, C. *et al.* Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. *N. Engl. J. Med.* **362**, 875–885 (2010).
- Shustik, J. *et al.* Correlations between BCL6 rearrangement and outcome in patients with diffuse large B-cell lymphoma treated with CHOP or R-CHOP. *Haematologica* **95**, 96–101 (2010).
- Martinez-Ramirez, A. *et al.* Simultaneous detection of the immunophenotypic markers and genetic aberrations on routinely processed paraffin sections of lymphoma samples by means of the FICTION technique. *Leukemia* **18**, 348–353 (2004).
- Klein, U. *et al.* Transcriptional analysis of the B cell germinal center reaction. *Proc. Natl Acad. Sci. USA* **100**, 2639–2644 (2003).

The CatSper channel mediates progesterone-induced Ca^{2+} influx in human sperm

Timo Strücker^{1*}, Normann Goodwin^{1*}, Christoph Brenker^{1*}, Nachiket D. Kashikar^{1†}, Ingo Weyand², Reinhard Seifert¹ & U. Benjamin Kaupp¹

In the oviduct, cumulus cells that surround the oocyte release progesterone. In human sperm, progesterone stimulates a Ca^{2+} increase by a non-genomic mechanism^{1–3}. The Ca^{2+} signal has been proposed to control chemotaxis, hyperactivation and acrosomal exocytosis of sperm^{4–8}. However, the underlying signalling mechanism has remained mysterious. Here we show that progesterone activates the sperm-specific, pH-sensitive CatSper Ca^{2+} channel^{9–11}. We found that both progesterone and alkaline pH stimulate a rapid Ca^{2+} influx with almost no latency, incompatible with a signalling pathway involving metabotropic receptors and second messengers. The Ca^{2+} signals evoked by alkaline pH and progesterone are inhibited by the Ca_v channel blockers NNC 55-0396 and mibefradil. Patch-clamp recordings from sperm reveal an alkaline-activated current carried by mono- and divalent ions that exhibits all the hallmarks of sperm-specific CatSper Ca^{2+} channels^{10,11}. Progesterone substantially enhances the CatSper current. The alkaline- and progesterone-activated CatSper current is inhibited by both drugs. Our results resolve a long-standing controversy over the non-genomic progesterone signalling. In human sperm, either the CatSper channel itself or an associated protein serves as the non-genomic progesterone receptor. The identification of CatSper channel blockers will greatly facilitate the study of Ca^{2+} signalling in sperm and help to define further the physiological role of progesterone and CatSper.

Progesterone-evoked Ca^{2+} influx in human sperm does not involve classical regulation of transcription by nuclear receptors. Various candidate membrane receptors for progesterone have emerged, including a novel G-protein-coupled-type progestin receptor (mPR) and a single-pass receptor (progesterone-receptor membrane component, PGRMC)¹². The pathway downstream of the non-genomic progesterone receptor has been proposed to involve cAMP and cGMP, protein kinase A and G, Ca^{2+} release from intracellular stores, store-operated Ca^{2+} channels and cGMP-activated channels^{1,8,13}.

Here we study progesterone action in human sperm using optical and electrophysiological techniques. In sperm loaded with the Ca^{2+} indicator Fluo-4, progesterone elicited a change in intracellular Ca^{2+} concentration ($[\text{Ca}^{2+}]_i$). The biphasic signal^{3,14} recorded in a plate reader consisted of a rapid Ca^{2+} transient followed by a slower, sustained Ca^{2+} elevation (Fig. 1a). The dose-response relation of the transient and the sustained components (Fig. 1b) yielded constants of half-maximal activation ($K_{1/2}$) of 42 ± 15 nM and 91 ± 31 nM (six experiments), respectively.

We tested whether the progesterone-induced Ca^{2+} signal is different in capacitated and non-capacitated sperm. Under the capacitating conditions used here, the amplitude and time course of Ca^{2+} signals were overall unaffected by incubation conditions that promote sperm capacitation (Supplementary Fig. 1a). However, under capacitating conditions, the $K_{1/2}$ values (10 ± 10 nM for the transient and

24 ± 35 nM for the sustained component, six experiments) were shifted to lower concentrations, indicating that capacitation renders sperm more sensitive to progesterone (Supplementary Fig. 1b).

Signalling involving activation of metabotropic receptors is usually characterized by a latency of the cellular response, for example chemosensory Ca^{2+} signalling in sea urchin sperm or G-protein-coupled receptor signalling in vertebrate olfactory neurons^{15–17}. In contrast, insect olfactory neurons, which use ionotropic chemoreceptors, respond with almost no latency¹⁷. To test whether progesterone activates a G-protein-coupled receptor pathway, we recorded Ca^{2+} signals using the kinetic stopped-flow technique (Fig. 1c). The progesterone-induced Ca^{2+} signals recorded in a plate reader (Fig. 1a, b) and in the stopped-flow apparatus (Fig. 1c, d) were similar. The Ca^{2+} response at first saturated at about 1 μM and then continued to rise at concentrations greater than 2 μM (Fig. 1c, d), suggesting that progesterone binds to sites of high and low affinity. The $K_{1/2}$ of the high-affinity site was 42.3 ± 9.3 nM (three experiments) (Fig. 1d). Owing to the poor aqueous solubility of progesterone, we could not determine a $K_{1/2}$ value for the low-affinity site. At all progesterone concentrations, $[\text{Ca}^{2+}]_i$ rose within the time resolution of the stopped-flow apparatus (36 ms) (Fig. 1e). The instantaneous rise of $[\text{Ca}^{2+}]_i$ argues for an extracellular action of progesterone on a membrane protein, consistent with the observation that impermeable progesterone derivatives also stimulate a Ca^{2+} response¹⁸.

We investigated which signal component rests upon Ca^{2+} entry. Sperm were mixed in the stopped-flow apparatus with a Ca^{2+} -free medium containing progesterone and different concentrations of the Ca^{2+} chelator BAPTA (Fig. 1f): thereby, within milliseconds, a well-defined external Ca^{2+} concentration ($[\text{Ca}^{2+}]_o$) is adjusted, which prevents depletion of intracellular Ca^{2+} stores before progesterone stimulation. The amplitude of the transient Ca^{2+} signal evoked by 1 μM progesterone strongly depended on $[\text{Ca}^{2+}]_o$ (Fig. 1f). At 28 nM $[\text{Ca}^{2+}]_o$, even high progesterone concentrations (10 μM) no longer produced a Ca^{2+} response (Supplementary Fig. 2a). Moreover, at such low $[\text{Ca}^{2+}]_o$, the slow sustained component was also abolished (Supplementary Fig. 2b). We conclude that both transient and sustained Ca^{2+} signals result from Ca^{2+} entry. Alternatively, the sustained component might reflect Ca^{2+} release from intracellular stores involving inositol trisphosphate (IP_3), a mechanism that is also triggered by Ca^{2+} influx (' Ca^{2+} -induced Ca^{2+} release')^{1,2,19,20}.

The instantaneous rise of $[\text{Ca}^{2+}]_i$ is difficult to reconcile with a mechanism involving metabotropic receptors or the synthesis of second messengers. However, progesterone has been reported to elevate cAMP levels^{8,21} and cAMP to elevate $[\text{Ca}^{2+}]_i$ (refs 9, 22). We failed to reproduce these results. Progesterone did not elevate cAMP levels (Fig. 2a). Moreover, colforsin, which activates all membrane adenylyl cyclases²³, did not affect cAMP levels, arguing altogether against canonical G-protein-coupled receptor-operated cAMP signalling in human sperm. By contrast, HCO_3^- , which stimulates

¹Center of Advanced European Studies and Research, Abteilung Molekulare Neurosensorik, Ludwig-Erhard-Allee 2, 53175 Bonn, Germany. ²Institute of Complex Systems – Cellular Biophysics (ICS-4), Forschungszentrum Jülich, 52425 Jülich, Germany. [†]Present address: Max Planck Institute for Biophysical Chemistry, Max Planck Research Group Sleep and Waking, Am Faßberg 11, D-37077 Göttingen, Germany.

*These authors contributed equally to this work.

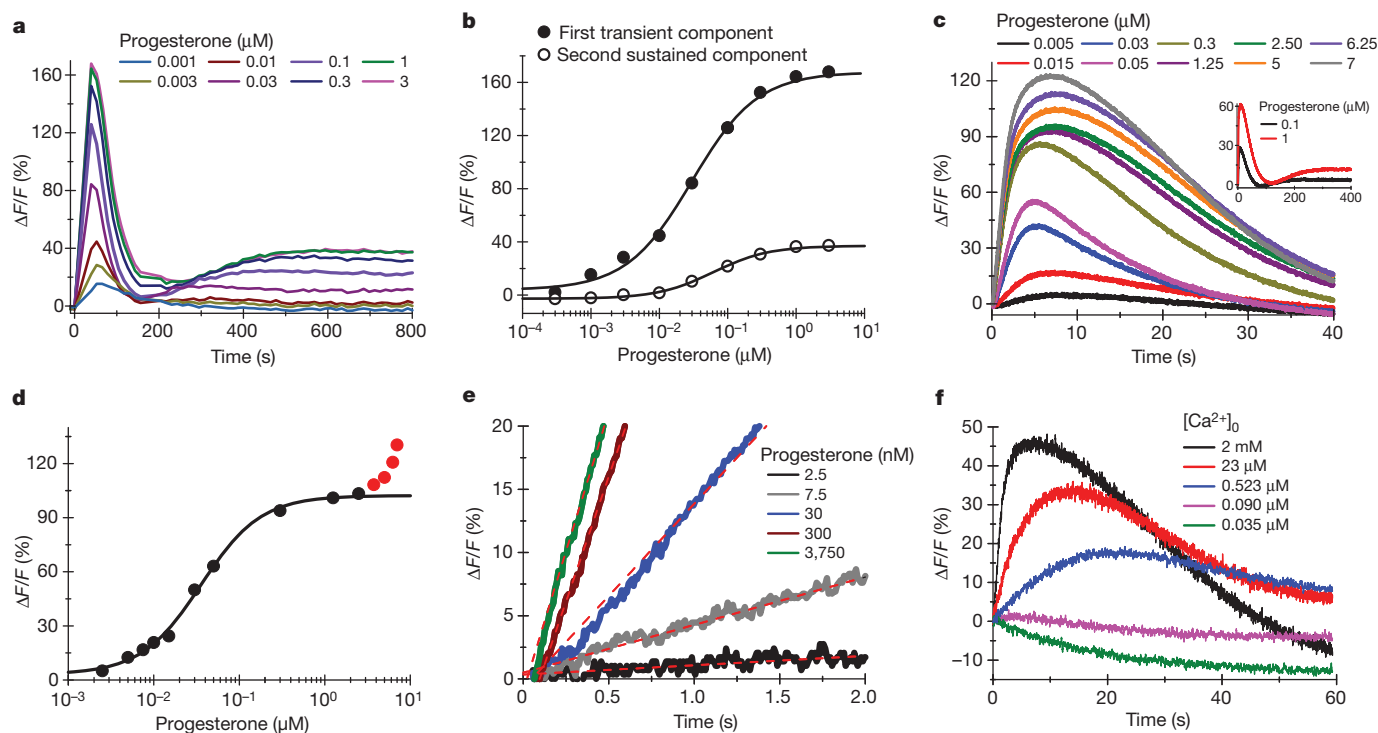


Figure 1 | Progesterone-induced Ca^{2+} signals in human sperm. **a**, Signals measured in a fluorescence plate reader. **b**, Dose–response relation of the two kinetic signal components from **a** ($K_{1/2}$: first component 33 nM; second component 68 nM). **c**, Signals measured in a stopped-flow apparatus. Inset: signals recorded for 400 s. **d**, Dose–response relation of signals from **c** ($K_{1/2}$:

35 nM). Red dots: continuous rise of $[\text{Ca}^{2+}]_i$ at progesterone concentrations greater than 2 μM . **e**, First 2 s of Ca^{2+} signals. Red lines: linear fit to their initial slope. **f**, Ca^{2+} signals evoked by 1 μM progesterone at different external Ca^{2+} concentrations $[\text{Ca}^{2+}]_o$ in the stopped-flow apparatus.

a soluble adenylyl cyclase²³, enhanced cAMP levels by fourfold. Isobutylmethylxanthine (IBMX), which inhibits phosphodiesterases, increased cAMP levels by 4.9-fold. A combination of HCO_3^- and IBMX augmented cAMP levels by 10.5-fold: that is, beyond values reached by each substance alone (Fig. 2a). Importantly, a rise of cAMP either by photolysis of caged cAMP or stimulation by HCO_3^- did not evoke Ca^{2+} increase (Fig. 2b). If anything, HCO_3^- caused a small decrease of $[\text{Ca}^{2+}]_i$. In mouse sperm also, HCO_3^- does not stimulate a Ca^{2+} response²⁴. Therefore, cAMP signalling is not directly involved in the rapid Ca^{2+} response to progesterone.

The principal voltage-gated Ca^{2+} channel (Ca_v) in epididymal mouse and ejaculated human sperm is CatSper^{10,11}, which is activated at alkaline intracellular pH (refs 10, 11). We reasoned that CatSper

mediates progesterone-induced Ca^{2+} entry. We tested several inhibitors of Ca_v channels for their ability to suppress progesterone- and alkaline-evoked Ca^{2+} responses. Only the T-type channel blockers NNC 55-0396 and mibefradil significantly impaired the progesterone-induced Ca^{2+} response (Fig. 3a, b). In addition, both compounds significantly impaired the alkaline-induced Ca^{2+} signals evoked by NH_4Cl (10 mM) (Fig. 3c–e). However, we emphasize that, in Ca^{2+} fluorimetry, the pharmacology of these drugs is complex: at concentrations greater than 10 μM NNC 55-0396 and greater than 40 μM mibefradil, the drugs evoked Ca^{2+} responses themselves (Supplementary Fig. 3a, b). This prevented us testing whether higher drug concentrations completely inhibited the progesterone- and alkaline-evoked Ca^{2+} signals. Progesterone does not stimulate Ca^{2+} entry by alkalization; in contrast to

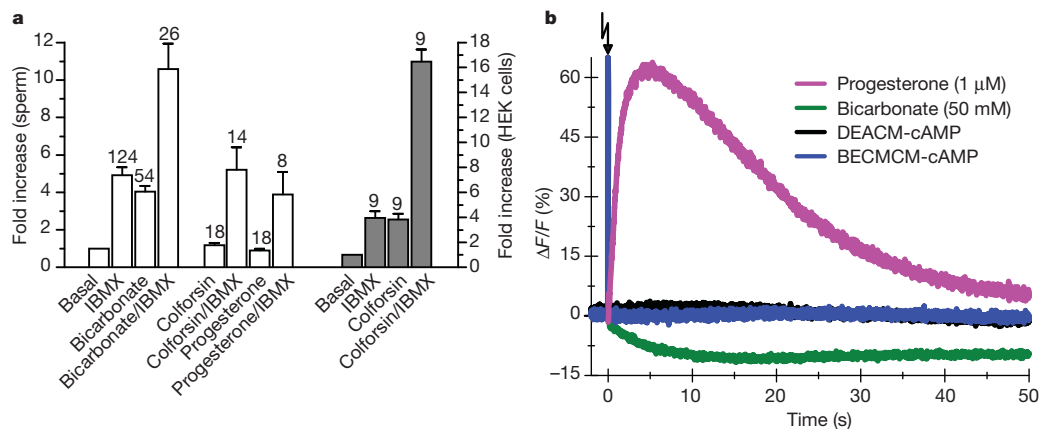


Figure 2 | Progesterone does not activate a cAMP–signalling pathway. **a**, Changes in total cAMP concentration in human sperm (white) and HEK 293 cells (grey) evoked by 0.5 mM IBMX, 50 mM bicarbonate, 50 μM colforsin, 1 μM progesterone (mean \pm s.e.m.; number of measurements above the bars, from 3 to 26 experiments). At rest, 10^6 human sperm contain

0.087 ± 0.002 pmol cAMP (mean \pm s.e.m., 290 measurements from 26 experiments), corresponding to approximately 2.5 μM cAMP (sperm volume ~ 35 fL). **b**, Ca^{2+} signals produced by photolysis of DEACM-cAMP and BECMCM-cAMP, and by bicarbonate and progesterone. Experiments were done in a stopped-flow apparatus.

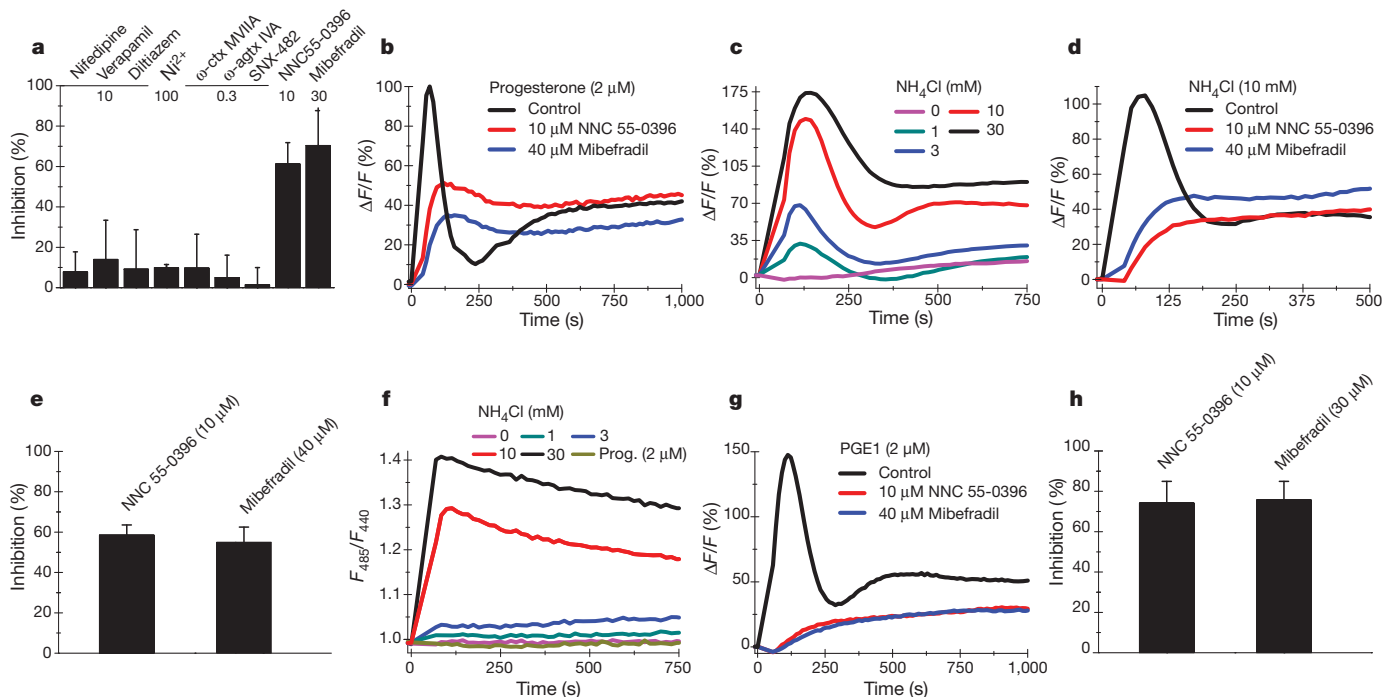


Figure 3 | Pharmacology of progesterone-induced Ca^{2+} signals. **a**, Relative inhibition of progesterone-evoked ($2 \mu\text{M}$) Ca^{2+} signals by Ca_v channel blockers (three to five experiments, inhibitor concentrations in μM). **b**, Progesterone-evoked Ca^{2+} signals are inhibited by NNC 55-0396 or mibefradil. **c**, NH_4Cl -evoked Ca^{2+} signals are inhibited by NNC 55-0396 or mibefradil. **d**, NH_4Cl responses are inhibited by NNC 55-0396 or mibefradil. **e**, Relative inhibition of NH_4Cl -induced Ca^{2+} signals by NNC 55-0396 or mibefradil (four experiments). **f**, NH_4Cl -evoked changes in pH_i . Progesterone did not change pH_i . **g**, PGE1-evoked Ca^{2+} signals inhibited by NNC 55-0396 or mibefradil. **h**, Relative inhibition of PGE1-evoked Ca^{2+} signals by NNC 55-0396 and mibefradil (four experiments). Experiments were done in a fluorescence plate reader; data are given as mean \pm s.d.

NH_4Cl , progesterone ($2 \mu\text{M}$) did not change pH_i (Fig. 3f). Taken together, these results suggest that progesterone activates CatSper.

Prostaglandin E1 (PGE1) and progesterone produce similar biphasic Ca^{2+} signals²⁵ (Figs 1a and 3g and Supplementary Fig. 4). The pharmacology of the PGE1 response is incompatible with known prostanoïd receptors²⁵. The PGE1-induced Ca^{2+} signal is also suppressed by NNC 55-0396 and mibefradil, indicating that PGE1 also activates CatSper (Fig. 3g, h). Sequential application of 17-OH-progesterone and progesterone, but not of PGE1 and progesterone²⁵, leads to cross-desensitization (Supplementary Fig. 4), suggesting that progesterone and PGE1 use distinct binding sites or signalling mechanisms.

We studied membrane currents in mature non-capacitated sperm by whole-cell patch-clamp recordings. In standard extracellular solution containing Ca^{2+} and Mg^{2+} , steps in membrane voltage V from 0 mV to ± 80 mV produced hardly any currents¹¹ (Fig. 4a). When switching to divalent-free solution ($\text{NaDVF} + 0 \text{ Prog}$; Fig. 4a), monovalent currents carried by CatSper channels appeared^{10,11}. Progesterone evoked a dose-dependent increase of monovalent current (Fig. 4a) with a $K_{1/2}$ of $66 \pm 19 \text{ nM}$ (five experiments) (Fig. 4b). Mean inward currents at -80 mV were $-59.4 \pm 26.8 \text{ pA}$, range -30.3 to -102.9 pA (0 progesterone) and $-180.8 \pm 73.5 \text{ pA}$, range -126.9 to -264.5 pA ($10 \mu\text{M}$ progesterone). Remarkably, the reversal potential (V_{rev}) of the monovalent current ($35 \pm 5 \text{ mV}$; five experiments) did not change during progesterone stimulation (Fig. 4c), indicating that V_{rev} of progesterone-induced and CatSper currents is similar. RU486, an antagonist for nuclear progesterone receptors, did not inhibit the progesterone-induced currents (Supplementary Fig. 5). Stimulation with NH_4Cl (10 mM), or progesterone ($1 \mu\text{M}$), or both, potentiated monovalent currents, again with no change in V_{rev} (Fig. 4d). Moreover, PGE1 also activated a monovalent current with a V_{rev} similar to that of CatSper currents (Supplementary Fig. 6a). Most importantly, mibefradil ($30 \mu\text{M}$) completely blocked the monovalent current activated by NH_4Cl , progesterone (Fig. 4d) and PGE1 (Supplementary Fig. 6b). This suggests that progesterone- and PGE1-induced currents are carried by CatSper.

Under physiological conditions, CatSper primarily carries divalent currents. Therefore, we studied the voltage dependence, pH sensitivity and pharmacology of Ba^{2+} tail currents elicited after repolarization from various test voltages (Fig. 4e, g). At pH_i 6, almost all CatSper channels are closed. At pH_i 6, progesterone ($1 \mu\text{M}$) or NH_4Cl (10 mM), however, induced sizeable Ba^{2+} tail currents (Fig. 4e). Simultaneous application of NH_4Cl and progesterone enhanced tail currents in a non-additive manner (Fig. 4e): the current stimulated by progesterone/ NH_4Cl was three times larger (2.96 ± 0.7 ; five experiments at $+70 \text{ mV}$) than the superposition of currents produced by each substance alone. Alkalization and progesterone shifted the activation curve to less positive voltages (Fig. 4f). The voltage shift evoked by progesterone/ NH_4Cl was larger than that caused by progesterone or NH_4Cl alone (Fig. 4f). At pH_i 8, a significant fraction of CatSper channels was activated (Fig. 4g, h); and progesterone ($1 \mu\text{M}$) further enhanced the Ba^{2+} tail currents (Fig. 4g, h). NNC 55-0396 ($10 \mu\text{M}$) inhibited these tail currents (Fig. 4h). In summary, the similar V_{rev} , the non-additive action of pH_i and progesterone, and the blockage by NNC 55-0396 and mibefradil demonstrate that alkaline pH_i , progesterone and PGE1 activate CatSper. We recorded progesterone-evoked CatSper currents using pipette solutions without ATP and GTP, arguing against an activation mechanism involving second messengers, phosphorylation or G proteins.

Here we show that, in human sperm, progesterone activates CatSper either by binding to the channel itself or to an associated protein. Based on electrophysiological analysis of progesterone-induced currents in sperm and isolated flagella, Lishko *et al.*²⁶ come to a similar conclusion. In mice, CatSper controls hyperactivation, a mode of sperm motility required for fertilization. Considering the function of progesterone as a chemoattractant^{6,8}, we suggest that CatSper serves in human sperm as a Ca^{2+} channel for both hyperactivation and chemotaxis. CatSper channels are also expressed in other mammalian species. Rabbit, human and bovine follicular factors display chemotactic activity across species, arguing that mammals use a common or similar chemoattractant²⁷. Further studies are required to examine whether other mammals also use progesterone as a chemoattractant and whether in these

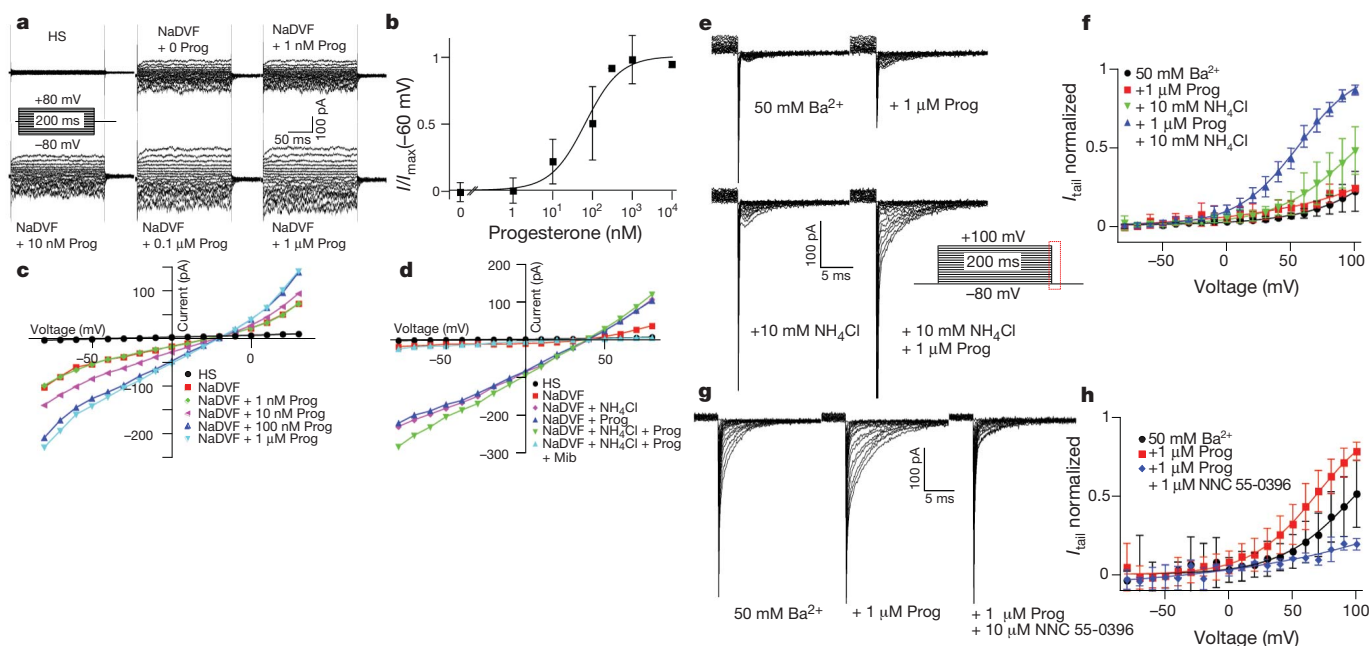


Figure 4 | Electrophysiological characterization of whole-cell CatSper currents from human sperm cells. **a**, Control currents in standard extracellular solution (HS) and monovalent currents in divalent-free Na^+ -based bath solution (NaDVF), in the absence of intracellular divalent ions. Currents were recorded at pH_i 7.3 after extracellular addition of different progesterone concentrations. Voltage was stepped from -80 mV to $+80$ mV in increments of 10 mV. **b**, Dose–response relation of the mean current at -60 mV ($K_{1/2}$: 66 ± 19 nM, Hill coefficient 1 ± 0.2 ; mean \pm s.d. of five experiments). **c**, Current–voltage relation of monovalent currents at different extracellular progesterone concentrations. **d**, Activation of monovalent CatSper currents at pH_i 7.3 by $1 \mu\text{M}$ extracellular progesterone, or intracellular

alkalization with 10 mM NH_4Cl , or both. Currents were completely blocked by $30 \mu\text{M}$ mibefradil. **e**, Ba^{2+} tail currents at -80 mV after voltage steps from -80 mV to $+100$ mV in increments of 10 mV at pH_i 6. The red box in the voltage protocol indicates the part of the current traces shown. **f**, Ba^{2+} tail currents at pH_i 6; mean \pm s.d. of four or five experiments; currents were normalized to the interpolated maximum current evoked by simultaneous application of progesterone and NH_4Cl . **g**, Ba^{2+} tail currents at -80 mV after voltage steps from -80 mV to $+100$ mV in increments of 10 mV at pH_i 8. Same protocol used as in **e**. **h**, Ba^{2+} tail currents at pH_i 8; mean \pm s.d. of two to six experiments; currents were normalized to the interpolated maximum current evoked by $1 \mu\text{M}$ progesterone.

species progesterone also activates CatSper. However, at least in mice, progesterone does not activate CatSper (Lishko *et al.*²⁶). Moreover, many substances as diverse as odorants and prostaglandins evoke Ca^{2+} responses in human sperm^{22,25,28}, and several chemicals or factors have been proposed to attract sperm^{4,28}. Chemotaxis in the female genital tract might be a multistep process that involves distinct chemoattractants and even thermotaxis⁴. Future work needs to show whether other substances that activate CatSper serve as chemoattractants, and whether substances that reportedly display chemotactic activity also open CatSper.

We propose that CatSper represents a ligand-gated Ca^{2+} channel that can be activated by lipophilic compounds. However, the binding site for progesterone has yet to be determined. We cannot exclude the possibility that progesterone binds to proteins such as the mPRs and/or PGRMCs that might associate with CatSper and convey the binding event to the channel. Such a mechanism, where distinct proteins form a receptor/channel complex, exists in insect olfactory neurons^{17,29,30}. Alternatively, the accessory subunits CatSper- β or CatSper-G might host the binding site. The availability of CatSper blockers provides a powerful tool to establish unequivocally progesterone as a chemoattractant and CatSper as a chemotaxis transduction channel, but also to interfere pharmacologically with fertilization.

METHODS SUMMARY

Sperm preparation. Sperm were purified by a ‘swim-up’ procedure in human tubular fluid (HTF) medium. Sperm were washed and resuspended in HTF⁺⁺ containing 4 mM NaHCO_3 and 3 mg ml^{-1} human serum albumin (Irvine Scientific). Before experiments, sperm were incubated for at least 1 h in HTF⁺⁺ at 37°C . Under these conditions, the sperm are non-capacitated. For capacitation, sperm were incubated in 25 mM NaHCO_3 and 3 mg ml^{-1} human serum albumin for at least 2 h. Unless otherwise specified, experiments were done with non-capacitated sperm.

Measurement of changes in intracellular Ca^{2+} and pH. Changes in $[\text{Ca}^{2+}]_i$ and pH_i were measured in sperm loaded with Fluo-4 and BCECF (Molecular Probes), respectively, in 384 multi-well plates in a fluorescence plate reader (Fluostar Omega, BMG Labtech) at 29°C and in a rapid-mixing device in the stopped-flow mode (SFM400, Bio-Logic) at 37°C . After loading, sperm were washed and resuspended in HTF⁺⁺ containing 4 mM NaHCO_3 but no serum albumin, unless otherwise specified. Ligands or inhibitors were dissolved in HTF⁺⁺ as well.

Caged compounds and photolysis. We used the caged cyclic nucleotides BECMCM-caged cAMP and DEACM-caged cAMP. The caged compound was photolysed by an ultraviolet flash from a xenon flash lamp (DP-10, Rapp OptoElectronic).

Determination of cAMP content. Sperm in HTF⁺⁺ were mixed 1:1 (v/v) with substances dissolved in HTF⁺⁺. The cAMP content was determined by radioimmunoassays (^{125}I -labelled cAMP; IBL).

Patch-clamp recordings. We recorded from sperm in the whole-cell configuration. Seals between pipette and spermatozoa were formed either at the cytoplasmic droplet or in the neck region in standard extracellular solution (HS) containing divalent cations. Monovalent currents were recorded in a sodium-based divalent-free solution (NaDVF). Bath solutions for recording divalent CatSper currents contained 50 mM Ba^{2+} . Different pipette solutions were used depending on whether monovalent or divalent currents were recorded.

Data analysis. Data are given as mean \pm s.d. (number of experiments) if not otherwise stated.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 July; accepted 17 December 2010.

- Publicover, S., Harper, C. V. & Barratt, C. $[\text{Ca}^{2+}]_i$ signalling in sperm – making the most of what you’ve got. *Nature Cell Biol.* **9**, 235–242 (2007).
- Thomas, P. & Meizel, S. Phosphatidylinositol 4,5-bisphosphate hydrolysis in human sperm stimulated with follicular fluid or progesterone is dependent upon Ca^{2+} influx. *Biochem. J.* **264**, 539–546 (1989).

3. Blackmore, P. F., Beebe, S. J., Danforth, D. R. & Alexander, N. Progesterone and 17 α -hydroxyprogesterone. Novel stimulators of calcium influx in human sperm. *J. Biol. Chem.* **265**, 1376–1380 (1990).
4. Eisenbach, M. & Giojalas, L. C. Sperm guidance in mammals – an unpaved road to the egg. *Nature Rev. Mol. Cell Biol.* **7**, 276–285 (2006).
5. Harper, C. V., Kirkman-Brown, J. C., Barratt, C. L. R. & Publicover, S. J. Encoding of progesterone stimulus intensity by intracellular $[Ca^{2+}]$ ($[Ca^{2+}]_i$) in human spermatozoa. *Biochem. J.* **373**, 407–417 (2003).
6. Oren-Benaroya, R., Orvieto, R., Gakamsky, A., Pinchasov, M. & Eisenbach, M. The sperm chemoattractant secreted from human cumulus cells is progesterone. *Hum. Reprod.* **23**, 2339–2345 (2008).
7. Suarez, S. S. Control of hyperactivation in sperm. *Hum. Reprod. Update* **14**, 647–657 (2008).
8. Teves, M. E. *et al.* Molecular mechanism for human sperm chemotaxis mediated by progesterone. *PLoS ONE* **4**, e8211 (2009).
9. Ren, D. *et al.* A sperm ion channel required for sperm motility and male fertility. *Nature* **413**, 603–609 (2001).
10. Kirichok, Y., Navarro, B. & Clapham, D. E. Whole-cell patch-clamp measurements of spermatozoa reveal an alkaline-activated Ca^{2+} channel. *Nature* **439**, 737–740 (2006).
11. Lishko, P. V., Botchkina, I. L., Fedorenko, A. & Kirichok, Y. Acid extrusion from human spermatozoa is mediated by flagellar voltage-gated proton channel. *Cell* **140**, 327–337 (2010).
12. Gellersen, B., Fernandes, M. S. & Brosens, J. J. Non-genomic progesterone actions in female reproduction. *Hum. Reprod. Update* **15**, 119–138 (2009).
13. Baldi, E. *et al.* Nongenomic activation of spermatozoa by steroid hormones: facts and fictions. *Mol. Cell. Endocrinol.* **308**, 39–46 (2009).
14. Kirkman-Brown, J. C., Bray, C., Stewart, P. M., Barratt, C. L. R. & Publicover, S. J. Biphasic elevation of $[Ca^{2+}]_i$ in individual human spermatozoa exposed to progesterone. *Dev. Biol.* **222**, 326–335 (2000).
15. Strücker, T. *et al.* A K^+ -selective cGMP-gated ion channel controls chemosensation of sperm. *Nature Cell Biol.* **8**, 1149–1154 (2006).
16. Kaupp, U. B. *et al.* The signal flow and motor response controlling chemotaxis of sea urchin sperm. *Nature Cell Biol.* **5**, 109–117 (2003).
17. Kaupp, U. B. Olfactory signalling in vertebrates and insects: differences and commonalities. *Nature Rev. Neurosci.* **11**, 188–200 (2010).
18. Blackmore, P. F., Neulen, J., Lattanzio, F. & Beebe, S. J. Cell surface-binding sites for progesterone mediate calcium uptake in human sperm. *J. Biol. Chem.* **266**, 18655–18659 (1991).
19. Bedu-Addo, K. *et al.* Mobilisation of stored calcium in the neck region of human sperm—a mechanism for regulation of flagellar activity. *Int. J. Dev. Biol.* **52**, 615–626 (2008).
20. Costello, S. *et al.* Ca^{2+} -stores in sperm: their identities and functions. *Reproduction* **138**, 425–437 (2009).
21. Parinaud, J. & Milhet, P. Progesterone induces Ca^{++} -dependent 3',5'-cyclic adenosine monophosphate increase in human sperm. *J. Clin. Endocrinol. Metab.* **81**, 1357–1360 (1996).
22. Spehr, M. *et al.* Identification of a testicular odorant receptor mediating human sperm chemotaxis. *Science* **299**, 2054–2058 (2003).
23. Kamenetsky, M. *et al.* Molecular details of cAMP generation in mammalian cells: a tale of two systems. *J. Mol. Biol.* **362**, 623–639 (2006).
24. Carlson, A. E., Hille, B. & Babcock, D. F. External Ca^{2+} acts upstream of adenylyl cyclase SACY in the bicarbonate signaled activation of sperm motility. *Dev. Biol.* **312**, 183–192 (2007).
25. Schaefer, M., Hofmann, T., Schultz, G. & Gudermand, T. A new prostaglandin E receptor mediates calcium influx and acrosome reaction in human spermatozoa. *Proc. Natl Acad. Sci. USA* **95**, 3008–3013 (1998).
26. Lishko, P. V., Botchkina, I. L. & Kirichok, Y. Progesterone activates the principal Ca^{2+} channel of human sperm. *Nature* doi:10.1038/nature09767 (this issue).
27. Sun, F. *et al.* Lack of species-specificity in mammalian sperm chemotaxis. *Dev. Biol.* **255**, 423–427 (2003).
28. Kaupp, U. B., Kashikar, N. D. & Weyand, I. Mechanisms of sperm chemotaxis. *Annu. Rev. Physiol.* **70**, 93–117 (2008).
29. Sato, K. *et al.* Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002–1006 (2008).
30. Wicher, D. *et al.* *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**, 1007–1011 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the German Research Foundation (SFB 645) and the International Helmholtz Research School on Biophysics and Soft Matter. We thank S. Stark for technical assistance, H. Krause and B. Kayser for preparing the manuscript, and C. Bernsdorff for preparing the figures. The caged compounds were kindly provided by V. Hagen.

Author Contributions T.S., N.G., C.B., N.K. and I.W. designed and performed experiments. R.S. designed experiments. T.S. and U.B.K. conceived the project. T.S. and U.B.K. wrote the manuscript. All authors revised and edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to U.B.K. (u.b.kaupp@caesar.de) or T.S. (timo.struenker@caesar.de).

METHODS

Sperm preparation. Human semen samples were obtained from healthy volunteers with their consent. Fresh ejaculates were allowed to liquefy at room temperature for 30–60 min. Sperm were purified by a 'swim-up' procedure. Liquefied semen (0.5–1 ml) was layered in a 50-ml falcon tube below 4 ml of HTF medium containing (in mM) 97.8 NaCl, 4.69 KCl, 0.2 MgSO₄, 0.37 KH₂PO₄, 2.04 CaCl₂, 0.33 sodium pyruvate, 21.4 sodium lactate, 2.78 glucose, and 21 HEPES, adjusted to pH 7.3–7.4 with NaOH. Motile sperm were allowed to swim-up into the HTF layer for 60–90 min at 37 °C and washed two times (700g, 20 min, room temperature (approximately 22 °C)). For cAMP determination, 'swim-up' sperm from several donors were pooled. For cAMP determination, samples were washed three times. The three washing steps are necessary to remove from seminal fluid an unknown factor that interferes with the cAMP assay and produces artificially high cAMP levels. Sperm number was determined in a Neubauer cell counter. Washed sperm were re-suspended in HTF⁺⁺ containing 4 mM NaHCO₃ and 3 mg ml⁻¹ human serum albumin (Irvine Scientific) at a density of 2×10^7 ml⁻¹, unless otherwise specified. Capacitation was assessed using the FITC-CD46 assay, which probes the complete acrosome reaction induced by a Ca²⁺ ionophore.

Measurement of changes in intracellular Ca²⁺ and pH. Changes in [Ca²⁺]_i and pH_i were measured in sperm loaded with Fluo-4 and BCECF (Molecular Probes), respectively, in 384 multi-well plates in a fluorescence plate reader (Fluostar Omega, BMG Labtech) at 29 °C and in a rapid-mixing device in the stopped-flow mode (SFM400, Bio-Logic) at 37 °C. Sperm were loaded with the fluorescent Ca²⁺ indicator Fluo-4 AM (10 µM) in the presence of Pluronic F-127 (0.05% v/v) at 37 °C for 45 min (ref. 31). To measure pH_i changes in the plate reader, sperm were loaded for 15–20 min with the fluorescent pH indicator BCECF (10 µM). After incubation, excess dye was removed by two centrifugation steps (700g, 10 min, room temperature (approximately 22 °C)). The pellet was re-suspended in the same volume of HTF containing 4 mM NaHCO₃ but no human serum albumin (HTF⁺) and equilibrated for 5 min at 37 °C. Each well was filled with 27 µl (1×10^7 sperm ml⁻¹) or 54 µl (5×10^6 sperm ml⁻¹) of the sperm suspension; the fluorescence was excited at 480 nm (Fluo-4) or alternating at 440 nm and 480 nm (BCECF, dual excitation). Fluorescence emission was recorded at 520 nm with bottom optics. Fluorescence was recorded before and after injection of 27 µl (1:1 dilution) or 6 µl (1:10 dilution) of solutions containing ligands, inhibitors or HTF⁺. The solutions were injected into the wells manually with an electronic multichannel pipette.

In the stopped-flow device, the sperm suspension was rapidly mixed (1:1 v/v; flow rate 1 ml s⁻¹) with HTF⁺ solutions containing different concentrations of ligand or other substances. At this flow rate, no physical damage of sperm was detected and their swimming behaviour appeared normal. Fluorescence excitation and recording were as described³¹.

Caged compounds and photolysis. The following caged cyclic nucleotides were used: BECMCM-caged cAMP and DEACM-caged cAMP^{15,16}. Sperm were loaded with 30 µM caged cyclic nucleotides for 30–45 min in the presence of Fluo-4 AM and Pluronic F-127. The ultraviolet flash was delivered to the cuvette by a liquid light guide and passed through a 295–395 (1 mm) band-pass filter (Rapp OptoElectronic).

Determination of cAMP content. The sperm suspension (1.25×10^6 cells) was mixed 1:1 (v/v) with different substances in HTF⁺⁺. The final concentration of substances was as follows (in mM): 50 HCO₃⁻, 0.5 IBMX; (in µM): 50 colforsin, 1 progesterone. The final concentration of solvent (0.5% DMSO) did not affect the resting cAMP level. After stimulation for 5 s to 25 min at 37 °C, the reaction was quenched with HClO₄ (1:3 (v/v); 0.5 M final concentration). Samples were neutralized by K₃PO₄ (0.24 M final concentration). The salt precipitate and cell debris were sedimented by centrifugation (15 min, 15,000g, 4 °C). The cAMP content in the supernatant (200 µl) was determined by radioimmunoassay that included an acetylation step for higher sensitivity (¹²⁵I-labelled cAMP; IBL). Calibration curves were obtained by serial dilutions of cAMP standards. The basal cAMP concentration was constant for up to 3 h. The cAMP increase produced by IBMX was largely identical for stimulation periods between 15 s and 300 s; pre-incubation of sperm with IBMX was for 2 min.

Patch-clamp recordings. We recorded from sperm in the whole-cell configuration. Seals between pipette and spermatozoa were formed either at the cytoplasmic droplet or in the neck region in standard extracellular solution (HS) containing (in mM) 135 NaCl, 5 KCl, 1 MgSO₄, 1 CaCl₂, 5 glucose, 1 sodium pyruvate, 10 lactic acid, 20 HEPES (pH 7.4 was adjusted with NaOH). Voltage pulses of 500–650 mV for 0.5 ms combined with light suction achieved transition into the whole-cell mode. Monovalent currents were recorded in divalent-free solution (NaDVF) (in mM): 140 NaCl, 40 HEPES and 1 EGTA (pH 7.4 was adjusted with NaOH); the pipette (10–15 MΩ) solution contained (in mM) 121 Cs-methanesulphonate, 65 HEPES, 4.5 EGTA, 4.5 CsCl (pH 7.3 was adjusted with CsOH); up to 20 mM glucose were added to the pipette solution to balance osmolarity. Pipette (15–25 MΩ) solution for recording of divalent currents contained (in mM) 165 N-methyl-D-glucamine (NMDG), 5 CsCl, 10 ethylene glycol tetraacetic acid (EGTA) and 10 MES or HEPES (pH 6.0 or pH 8.0 was adjusted with methanesulphonic acid (CH₃SO₃H)). Bath solution for recordings of divalent CatSper currents contained (in mM) 50 Ba(OH)₂, 90 NMDG, 20 HEPES (pH 7.4 was adjusted with methanesulphonic acid). NH₄Cl, progesterone, PGE1, NNC 55-0396 and mibefradil were added as indicated. Experiments were performed at 24 °C.

Data analysis. Data are given as mean ± s.d. (number of experiments) if not otherwise stated.

31. Kilic, F. *et al.* Caged progesterone: a new tool for studying rapid nongenomic actions of progesterone. *J. Am. Chem. Soc.* **131**, 4027–4030 (2009).

Progesterone activates the principal Ca^{2+} channel of human sperm

Polina V. Lishko¹, Inna L. Botchkina¹ & Yuriy Kirichok¹

Steroid hormone progesterone released by cumulus cells surrounding the egg is a potent stimulator of human spermatozoa. It attracts spermatozoa towards the egg and helps them penetrate the egg's protective vestments¹. Progesterone induces Ca^{2+} influx into spermatozoa^{1–3} and triggers multiple Ca^{2+} -dependent physiological responses essential for successful fertilization, such as sperm hyperactivation, acrosome reaction and chemotaxis towards the egg^{4–8}. As an ovarian hormone, progesterone acts by regulating gene expression through a well-characterized progesterone nuclear receptor⁹. However, the effect of progesterone upon transcriptionally silent spermatozoa remains unexplained and is believed to be mediated by a specialized, non-genomic membrane progesterone receptor^{5,10}. The identity of this non-genomic progesterone receptor and the mechanism by which it causes Ca^{2+} entry remain fundamental unresolved questions in human reproduction. Here we elucidate the mechanism of the non-genomic action of progesterone on human spermatozoa by identifying the Ca^{2+} channel activated by progesterone. By applying the patch-clamp technique to mature human spermatozoa, we found that nanomolar concentrations of progesterone dramatically potentiate CatSper, a pH-dependent Ca^{2+} channel of the sperm flagellum. We demonstrate that human CatSper is synergistically activated by elevation of intracellular pH and extracellular progesterone. Interestingly, human CatSper can be further potentiated by prostatic glandins, but apparently through a binding site other than that of progesterone. Because our experimental conditions did not support second messenger signalling, CatSper or a directly associated protein serves as the elusive non-genomic progesterone receptor of sperm. Given that the CatSper-associated progesterone receptor is sperm specific and structurally different from the genomic progesterone receptor, it represents a promising target for the development of a new class of non-hormonal contraceptives.

CatSper is a sperm-specific Ca^{2+} channel located in the principal piece of the flagellum^{11,12}. Weakly voltage-dependent, pH-sensitive CatSper is the only constitutively active Ca^{2+} conductance present in mouse and human spermatozoa as recorded using the whole-cell patch-clamp technique^{12,13}. Ca^{2+} influx through CatSper triggers sperm hyperactivation, a special high-amplitude asymmetrical flagellar beat required for penetration through viscous luminal fluids of the female reproductive tract and protective vestments of the egg^{14–16}. CatSper is also ideally positioned to control sperm chemotaxis, because 'chemotactic turns' that guide spermatozoa towards the egg depend on asymmetrical flagellar motion triggered by Ca^{2+} influx into the flagellum^{4,17}. Furthermore, CatSper-mediated Ca^{2+} influx into the flagellum leads to Ca^{2+} elevation even in the sperm head¹⁸ (probably by causing Ca^{2+} -dependent Ca^{2+} release from the intracellular store located in the sperm neck¹⁹) and thus can contribute to Ca^{2+} -dependent acrosome reaction²⁰. Because CatSper may cause all three Ca^{2+} -dependent responses triggered by progesterone (hyperactivation, chemotaxis and acrosome reaction), we tested the hypothesis that progesterone activates CatSper using the recently developed method for patch-clamping completely mature ejaculated human spermatozoa¹³.

Under normal physiological conditions, mouse and human CatSper channels are Ca^{2+} selective, but pass monovalent ions (Cs^+ or Na^+) under divalent-free conditions^{12,13}. Because monovalent CatSper currents are significantly larger, we studied CatSper currents under divalent-free conditions. The monovalent human CatSper current (I_{CatSper}) was overall smaller (Fig. 1a, blue) than mouse I_{CatSper} (Fig. 1c, blue), especially at negative membrane potentials (inward current). The virtual absence of human I_{CatSper} at the negative potentials normally found across the sperm plasma membrane was puzzling. Interestingly, addition of 500 nM progesterone to the bath solution dramatically increased the amplitude of human monovalent I_{CatSper} (Fig. 1a, red). Mouse monovalent I_{CatSper} did not increase after addition of 500 nM progesterone (Fig. 1c, red) or 10 μM progesterone (Supplementary Fig. 1).

Similar results were obtained with divalent I_{CatSper} . At intracellular pH 7.0, human Ba^{2+} I_{CatSper} was very small but increased dramatically after addition of 500 nM progesterone (Supplementary Fig. 2a). In contrast, mouse Ba^{2+} I_{CatSper} was easily detected at intracellular pH 7.0 and was not affected by progesterone (Supplementary Fig. 2b).

After potentiation with P, the amplitude of human I_{CatSper} remained stable. The potentiation was reversible and could be reproduced again on the same cell (Supplementary Fig. 3). Another major ovarian steroid hormone, oestradiol (17- β -oestradiol)^{9,10}, did not affect human CatSper (Supplementary Fig. 4).

We next conducted a series of experiments to confirm that the progesterone-activated current is indeed mediated by CatSper. Triturating human spermatozoa with a micropipette leads to the separation of a few spermatozoa into the head and the flagellum (containing midpiece and principal piece) at the neck region. Patch-clamp recording from isolated flagella detected monovalent CatSper currents that were strongly potentiated by progesterone and were indistinguishable from those recorded from whole human spermatozoa (Fig. 1b, compare with Fig. 1a). Mean amplitudes of currents recorded from whole human spermatozoa before and after stimulation with progesterone also closely matched those recorded from flagella (Fig. 2a), suggesting that similar to CatSper, the progesterone-activated conductance originates from the sperm flagellum.

Next, we discovered that human I_{CatSper} is potently blocked by an inhibitor of T-type voltage-gated Ca^{2+} channels, NNC55-0396 (ref. 21) (NNC, Fig. 2b, left panel), and tested whether NNC also inhibits the progesterone-activated current. Indeed, 2 μM NNC completely inhibited the whole-cell current in the presence of progesterone (Fig. 2b, right panel), confirming that the progesterone-activated current is carried by CatSper. Interestingly, because just 2 μM NNC completely suppressed human I_{CatSper} , we estimate that the affinity of NNC to CatSper is at least ten times greater than that to T-type voltage-gated Ca^{2+} channels (half-maximum inhibitory concentration (IC_{50}) \approx 7 μM for T-type channels)²¹. NNC blocked mouse CatSper in a similar range of concentrations (Supplementary Fig. 5). Furthermore, the Cs^+ I_{CatSper} was blocked by the same concentrations of Ca^{2+} as the progesterone-activated current, suggesting that the progesterone-activated channels

¹Department of Physiology, University of California San Francisco, UCSF Mail Code 2140, Genentech Hall Room N272F, 600 16th Street, San Francisco, California 94158, USA.

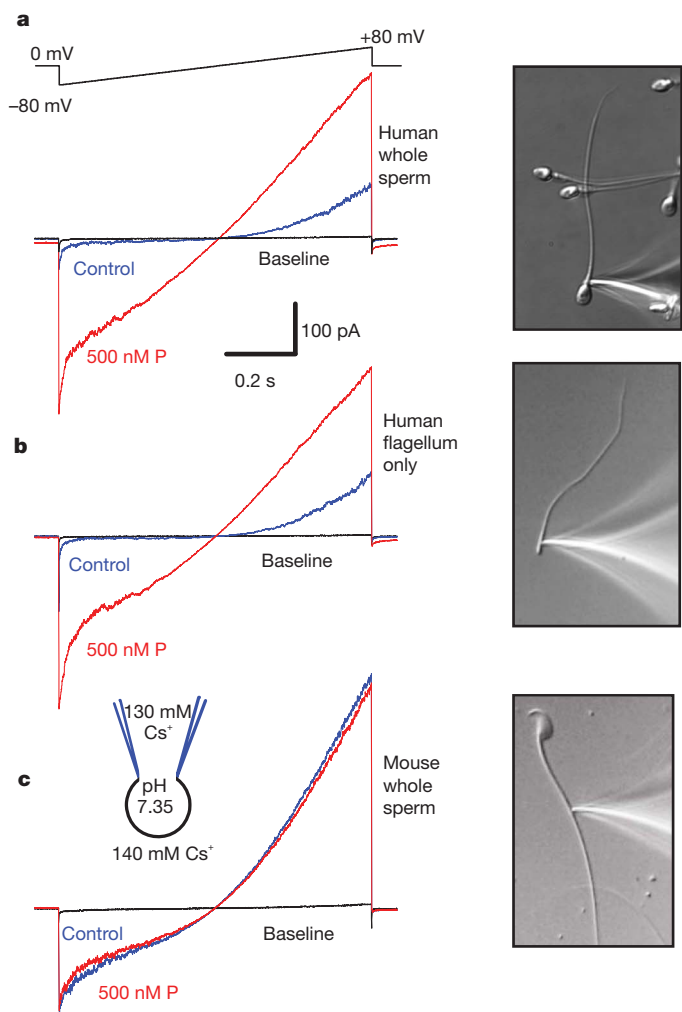


Figure 1 | Progesterone strongly potentiates human but not mouse CatSper. **a**, Representative monovalent whole-cell I_{CatSper} recorded from a human spermatozoon in the absence (blue) and presence (red) of 500 nM progesterone (P). Right, a human spermatozoon attached to the recording pipette. **b**, Representative whole-flagellum human I_{CatSper} . Right, sperm flagellum attached to the recording pipette. **c**, Representative whole-cell I_{CatSper} recorded from a mouse spermatozoon in the absence (blue) and presence (red) of 500 nM progesterone. Right, mouse spermatozoon attached to the recording pipette. Voltage protocol is shown in **a**. Baseline indicates currents recorded in HS solution.

contain a Ca^{2+} -binding site in the pore region that has the same affinity to Ca^{2+} as the binding site of CatSper (Fig. 2c and Supplementary Fig. 6).

Finally, we demonstrated that similar to CatSper, the progesterone-activated current is strongly potentiated by intracellular alkalization (Fig. 2d and Supplementary Fig. 7). Because the CatSper and progesterone-activated currents have identical properties, we conclude that the progesterone-activated current is mediated by CatSper. Identical slopes of the voltage-activation curves (with the same maximal conductance) for sperm Ca^{2+} currents observed with or without progesterone (Fig. 3c) also support this conclusion.

Progesterone potentiated CatSper with a half-maximum effective concentration (EC_{50}) of 7.7 ± 1.8 nM (Fig. 3a). The progesterone-binding site associated with CatSper is external, because a membrane-impermeable conjugate of progesterone and bovine serum albumin (BSA), P-3-BSA, activated I_{CatSper} similar to progesterone (Fig. 3b). Interestingly, progesterone modified with BSA at position C₁₁ (P-11-BSA) did not potentiate I_{CatSper} (Fig. 3b). A potent competitive antagonist of the nuclear progesterone receptor RU486 (mifepristone)²² did not inhibit potentiation of CatSper by progesterone (Supplementary Fig.

8), suggesting that the ligand-binding site of the CatSper-associated progesterone receptor is structurally different from that of the classic nuclear progesterone receptor.

CatSper activation by intercellular alkalization can be explained at least partially by shifting voltage-activation (G/V) curve to more negative, physiologically relevant potentials¹². Therefore we tested whether progesterone activates human CatSper through a similar mechanism. Indeed, at intracellular pH 7.4, addition of 500 nM progesterone shifted CatSper half-activation voltage ($V_{1/2}$) from +85 mV to +52 mV (Fig. 3c, black curves, and Supplementary Fig. 9a). On the other hand, the maximal amplitude of the CatSper tail currents did not change significantly after addition of 500 nM progesterone (control maximal tail current was $97 \pm 3\%$ of that with progesterone ($n = 7$); see Supplementary Fig. 9a). These results suggest that progesterone activates human CatSper primarily by shifting its voltage dependence towards more physiological, negative membrane potentials, while the single-channel conductance is not significantly affected.

Because *in vitro* capacitation (incubation of spermatozoa under conditions similar to those found in the Fallopian tubes) was shown to increase progesterone-induced Ca^{2+} transients in human spermatozoa²³, we investigated the effect of capacitation on the progesterone-induced shift in the CatSper G/V curve. Capacitation caused a 15 mV negative shift of the CatSper G/V curve, and slightly increased the amplitude of the negative shift induced by progesterone: addition of 500 nM progesterone to capacitated spermatozoa shifted CatSper half-activation voltage ($V_{1/2}$) from +70 mV to +30 mV (Fig. 3c, red curves, and Supplementary Fig. 9b). Again, no change in the maximal amplitude of the CatSper tail currents was observed after stimulation of capacitated spermatozoa with 500 nM progesterone: control maximal tail current was $99 \pm 4\%$ of that with progesterone ($n = 6$, Supplementary Fig. 9b). Thus capacitation enhances the effect of progesterone upon CatSper channel by providing additional negative shift of the CatSper G/V curve.

The experiments with CatSper voltage activation also revealed why, in the absence of progesterone, human I_{CatSper} is much smaller than mouse I_{CatSper} at negative membrane potentials (compare Fig. 1a and c). At intracellular pH 7.4, $V_{1/2}$ of human CatSper is approximately +85 mV (Fig. 3c), whereas under similar conditions ($\text{pH}_i = 7.5$), $V_{1/2}$ of mouse CatSper is approximately +11 mV (ref. 12). Moreover, the G/V curve of human CatSper is steeper (slope factor $k \approx 20$; Fig. 3c) than that of mouse CatSper ($k \approx 30$)¹². Slope factors for strongly voltage-sensitive ion channels are approximately 4. A more positive $V_{1/2}$ and a steeper slope of human CatSper G/V curve result in a smaller fraction of channels activated at negative membrane potentials compared with mouse CatSper. By inducing a negative shift in the G/V curve, progesterone helps human CatSper achieve a degree of activation at physiological potentials that is similar to mouse CatSper.

Prostaglandin E_1 (PGE_1) causes sperm intracellular Ca^{2+} transients similar in amplitude and waveform to those of progesterone-induced Ca^{2+} transients^{24–26}. Therefore we tested whether PGE_1 also potentiates CatSper. Indeed, addition of 500 nM PGE_1 induced strong potentiation of human monovalent I_{CatSper} (Fig. 4a), similar to that induced by progesterone. In contrast, mouse CatSper was not potentiated by PGE_1 , even at 10 μM (Supplementary Fig. 10). In human sperm, the large current recorded in the presence of PGE_1 was fully inhibited by 2 μM NNC (Fig. 4a), a potent CatSper channel blocker (Fig. 2b). As with progesterone-induced current, the PGE_1 -induced current originated from the flagellum (Fig. 4b, c) and was activated by intracellular alkalization induced by addition of 10 mM NH_4Cl (Fig. 4d and Supplementary Fig. 11). These findings confirm that, similar to progesterone, nanomolar concentrations of PGE_1 potentiate CatSper.

Although progesterone and PGE_1 activate the same channel, addition of a saturating concentration of PGE_1 (2 μM) after potentiation of I_{CatSper} with a saturating concentration of progesterone (2 μM) causes additional increase in the current amplitude and vice versa (Supplementary Fig. 12a, b). We used 2 μM as saturating concentration

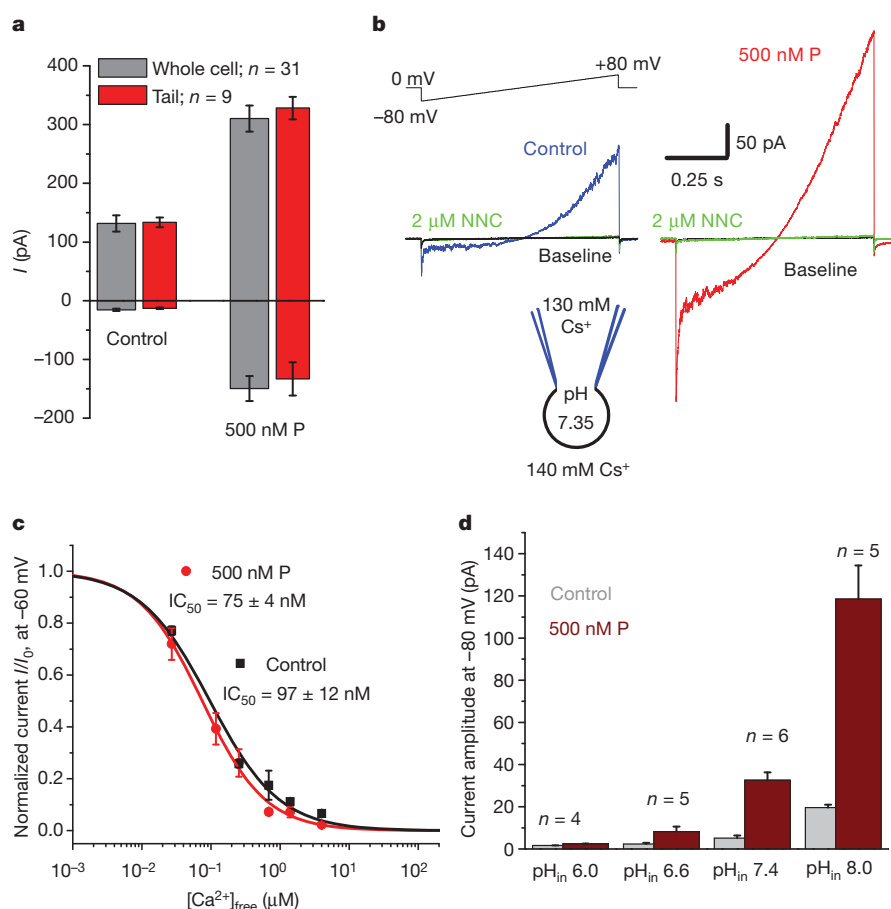


Figure 2 | Progesterone-activated current of human sperm is mediated by CatSper.

a, Averaged I_{CatSper} amplitudes recorded from a whole spermatozoon (grey) and a flagellum (red) in the absence and presence of progesterone (P). Current amplitudes at -80 mV (negative) and $+80$ mV (positive) are shown. **b**, Representative monovalent I_{CatSper} (left, blue) and the current in the presence of progesterone (right, red) were inhibited by $2 \mu\text{M}$ NNC (green). **c**, Dose-responses of Ca^{2+} inhibition of inward monovalent I_{CatSper} in the absence (black) and presence (red) of 500 nM progesterone. Amplitudes were recorded at -60 mV, $n = 10$. Representative current traces are shown in Supplementary Fig. 6. **d**, Amplitudes of inward monovalent I_{CatSper} at varying intracellular pH in the absence (grey) and presence (maroon) of 500 nM progesterone. Current amplitudes were measured at -80 mV. For representative traces, see Supplementary Fig. 7. Data are mean \pm s.e.m.

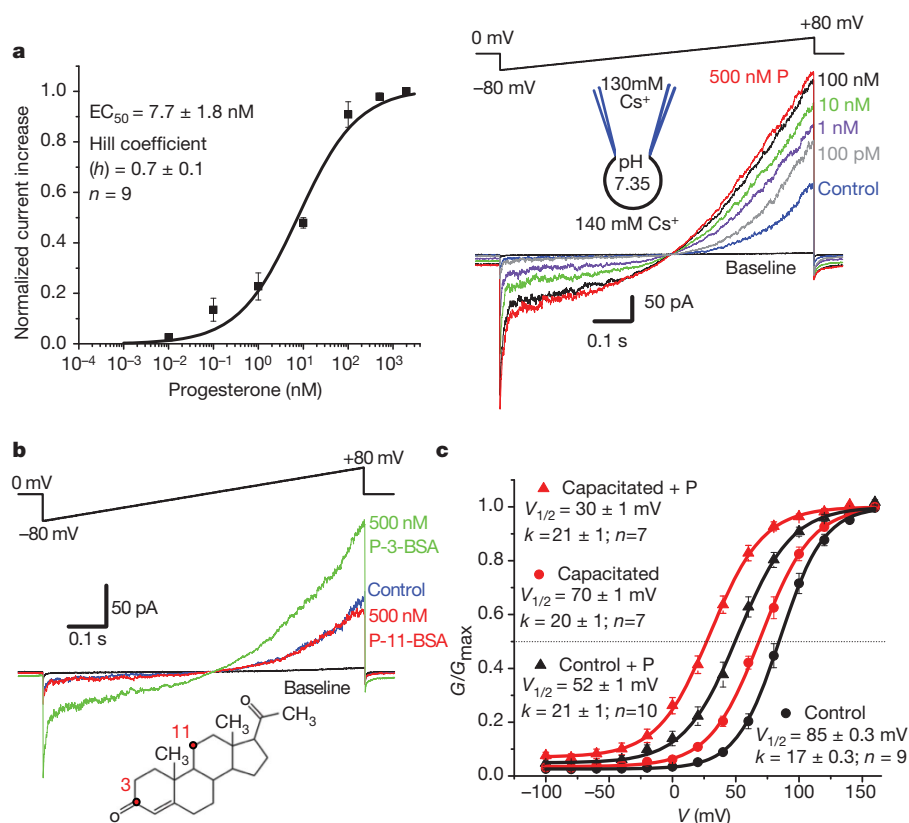


Figure 3 | Mechanism of CatSper potentiation by progesterone.

a, Dose dependence of potentiation of monovalent inward I_{CatSper} by progesterone (P; left panel) and representative monovalent whole-cell I_{CatSper} recorded from a human spermatozoon in the presence of varying concentrations of progesterone (right panel). Baseline indicates current recorded in HS solution. Current amplitudes were measured at -60 mV by averaging five to ten original current traces. **b**, Monovalent I_{CatSper} in control (blue) and in the presence of membrane-impermeable progesterone derivatives, P-3-BSA (green) and P-11-BSA (red). Below: chemical structure of progesterone with C_3 and C_{11} positions indicated. **c**, Voltage activation (G/V) curve of Ba^{2+} I_{CatSper} in the absence (circles) and presence (triangles) of 500 nM progesterone recorded in 10 mM Ba^{2+} bath solution. Black curves represent non-capacitated spermatozoa, whereas red curves represent capacitated spermatozoa. For representative traces, see Supplementary Fig. 9. Data are mean \pm s.e.m.

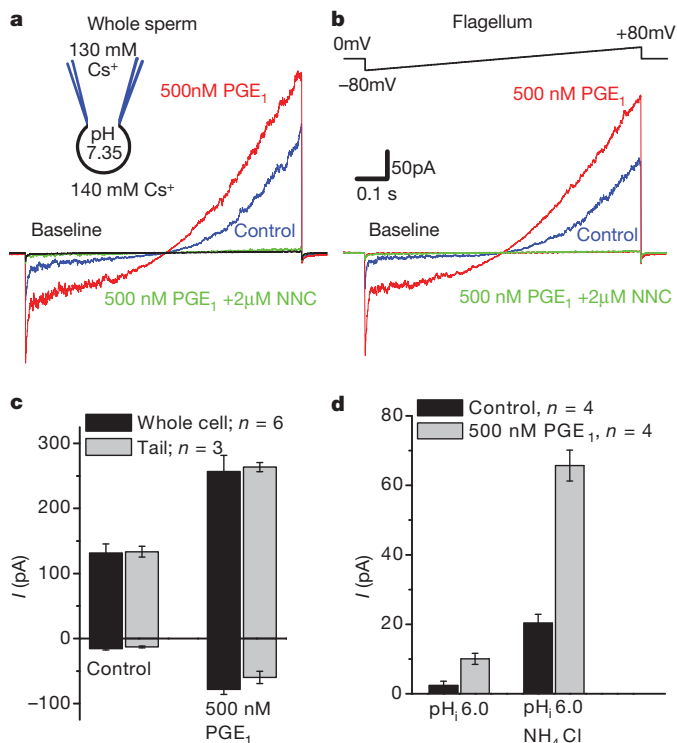


Figure 4 | Prostaglandin E₁ potentiates human CatSper. **a**, Representative whole-cell CatSper currents recorded from human spermatozoon in the absence (blue) and presence (red) of 500 nM PGE₁. Current with 500 nM PGE₁ was completely inhibited by 2 μM NNC (green). Baseline indicates current recorded in HS solution. **b**, Experiment from **a** reproduced with whole flagellum (principal piece and midpiece). **c**, Averaged *I*_{CatSper} amplitudes from a whole human spermatozoon (black) and a flagellum (grey) in the absence and presence of 500 nM PGE₁. Amplitudes of CatSper currents at -80 mV (negative) and +80 mV (positive) are shown. **d**, Potentiation of *I*_{CatSper} by 500 nM PGE₁ before and after induction of intracellular alkalinization with 10 mM NH₄Cl in the bath (initial intracellular pH = 6.0). Averaged *I*_{CatSper} amplitudes (at -80 mV) before (black) and after (grey) addition of 500 nM PGE₁. Data are mean ± s.e.m.

of PGE₁, because 10 μM PGE₁ did not induce any additional potentiation of *I*_{CatSper} (Supplementary Fig. 12c). These results suggest that the CatSper-associated binding sites for progesterone and PGE₁ are different and confirm a previous observation that progesterone and PGE₁ induce Ca²⁺ entry into human spermatozoa through different binding sites^{25,26}.

Prostaglandin F_{1α} (PGF_{1α}) activated CatSper as strongly as PGE₁. Prostaglandins E₂ (PGE₂) and A₁ (PGA₁) had weaker but significant effects, whereas prostaglandin D₂ (PGD₂) had almost no effect (Supplementary Fig. 13). The relative effects of all activators of human CatSper identified in this work followed the sequence progesterone > PGF_{1α} ≈ PGE₁ > PGA₁ > PGE₂ >> PGD₂ and are summarized in Supplementary Fig. 14.

The principal piece of the human sperm flagellum contains a voltage-gated proton channel H_v1 that may activate CatSper by causing intraflagellar alkalinization¹³. Because progesterone and PGE₁ had no effect on human sperm H_v1 current (Supplementary Fig. 15), we conclude that H_v1 is not involved in the sperm Ca²⁺ influx induced by progesterone or prostaglandins.

In summary, we demonstrate that human CatSper is strongly potentiated by progesterone and select prostaglandins, and that progesterone and PGE₁ apparently use different binding sites to activate CatSper. Because progesterone and prostaglandins potentiated CatSper in the absence of intracellular Ca²⁺, ATP or GTP, their effects are not mediated by G-proteins, protein kinases or second messengers. The simplest explanation of these results is that receptors for progesterone

and prostaglandins are located within the CatSper channel complex (consisting of CatSper1–4, CatSper-γ, CatSper-β²⁷). However, membrane progesterin receptors (mPR-α, -β, -γ, -δ and -ε) or progesterone receptor membrane component 1 (PGRMC1) that have been proposed to function as membrane progesterone receptors in different tissues^{28,29} may also serve as CatSper-associated progesterone receptors. Precise identification of the progesterone- and prostaglandin-binding sites is complicated by an inability to express functional CatSper channel complex in heterologous expression systems.

CatSper is present in different species including invertebrates, but orthologues of CatSper subunits have low identity (50% or less)^{27,30}, which probably reflects the fact that regulation of the main sperm Ca²⁺ entry pathway in species with different mechanisms of fertilization also differs. Indeed, in contrast to human CatSper, mouse CatSper is not activated by progesterone or prostaglandins.

The kinetic rapid-mixing fluorimetry used by Strünker *et al.*³¹ also demonstrates that CatSper is activated by progesterone and prostaglandins, without involvement of metabotropic receptors or second messengers.

METHODS SUMMARY

Gigohm seals between the patch pipette and mouse spermatozoa were formed at the cytoplasmic droplet. Human spermatozoa were patched either at the cytoplasmic droplet or, if the cytoplasmic droplet was inconspicuous, at the neck region. Seals were formed in HS solution comprising (in mM): 130 NaCl, 5 KCl, 1 MgSO₄, 2 CaCl₂, 5 glucose, 1 sodium pyruvate, 10 lactic acid, 20 HEPES, pH 7.4 adjusted with NaOH. Pipettes for whole-cell patch-clamp recordings of monovalent CatSper currents were filled with (in mM): 130 Cs-methanesulphonate (CsMeSO₃), 70 HEPES/MES, 3 EGTA, 2 EDTA, 0.5 TrisHCl, pH 6.0–8.0 adjusted with CsOH. Bath divalent-free solution for recording of monovalent CatSper currents contained the following (in mM): 140 CsMeSO₃, 40 HEPES/MES, 1 EDTA, pH 7.4 adjusted with CsOH. HS solution was used to record baseline current while measuring monovalent CatSper currents. Bath solution for recording monovalent CatSper currents at different bath [Ca²⁺]_{free} contained the following (in mM): 140 CsMeSO₃, 40 HEPES/MES, 1 EDTA, 1 BAPTA, 1 HEDTA, pH 7.4 adjusted with CsOH. CaCl₂ was added to this solution in accordance with WinMAXC version 2.05 (C. Patton, Stanford University) to obtain required free [Ca²⁺]. Pipettes for whole-cell patch-clamp recordings of Ba²⁺ CatSper currents were filled with the following (in mM): 145 NMDG, 100 HEPES, 10 BAPTA, 0.5 TrisHCl, pH 7.4 with HMeSO₃. Bath solution for recordings of Ba²⁺ currents contained the following (in mM): 0–50 Ba²⁺, 150–90 NMDG, 100 HEPES, pH 7.4 adjusted with HMeSO₃. Mg²⁺ bath solution used as a control (baseline) solution contained the following (in mM): 2 MgCl₂, 150 NMDG, 100 HEPES, pH 7.4 adjusted with HMeSO₃. Proton currents were recorded in divalent-free solution as described in Lishko *et al.*¹³. All electrophysiology experiments were performed at 24 °C. Data were analysed with Origin 7.0 and Clampfit 9.2. Statistical data were calculated as the mean ± s.e.m., and *n* indicates number of experiments.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 July; accepted 17 December 2010.

- Publicover, S., Harper, C. V. & Barratt, C. [Ca²⁺]_i signalling in sperm – making the most of what you've got. *Nature Cell Biol.* **9**, 235–242 (2007).
- Blackmore, P. F., Beebe, S. J., Danforth, D. R. & Alexander, N. Progesterone and 17 alpha-hydroxyprogesterone. Novel stimulators of calcium influx in human sperm. *J. Biol. Chem.* **265**, 1376–1380 (1990).
- Harper, C. V., Barratt, C. L. & Publicover, S. J. Stimulation of human spermatozoa with progesterone gradients to simulate approach to the oocyte. Induction of [Ca²⁺]_i oscillations and cyclical transitions in flagellar beating. *J. Biol. Chem.* **279**, 46315–46325 (2004).
- Eisenbach, M. & Giojalas, L. C. Sperm guidance in mammals – an unpaved road to the egg. *Nature Rev. Mol. Cell Biol.* **7**, 276–285 (2006).
- Revell, A., Massobrio, M. & Tesarik, J. Nongenomic actions of steroid hormones in reproductive tissues. *Endocr. Rev.* **19**, 3–17 (1998).
- Roldan, E. R., Murase, T. & Shi, Q. X. Exocytosis in spermatozoa in response to progesterone and zona pellucida. *Science* **266**, 1578–1581 (1994).
- Teves, M. E. *et al.* Progesterone at the picomolar range is a chemoattractant for mammalian spermatozoa. *Fertil. Steril.* **86**, 745–749 (2006).
- Uhler, M. L., Leung, A., Chan, S. Y. & Wang, C. Direct effects of progesterone and antiprogesterone on human sperm hyperactivated motility and acrosome reaction. *Fertil. Steril.* **58**, 1191–1198 (1992).
- Evans, R. M. The steroid and thyroid hormone receptor superfamily. *Science* **240**, 889–895 (1988).

10. Losel, R. & Wehling, M. Nongenomic actions of steroid hormones. *Nature Rev. Mol. Cell Biol.* **4**, 46–56 (2003).
11. Ren, D. *et al.* A sperm ion channel required for sperm motility and male fertility. *Nature* **413**, 603–609 (2001).
12. Kirichok, Y., Navarro, B. & Clapham, D. E. Whole-cell patch-clamp measurements of spermatozoa reveal an alkaline-activated Ca^{2+} channel. *Nature* **439**, 737–740 (2006).
13. Lishko, P. V., Botchkina, I. L., Fedorenko, A. & Kirichok, Y. Acid extrusion from human spermatozoa is mediated by flagellar voltage-gated proton channel. *Cell* **140**, 327–337 (2010).
14. Carlson, A. E. *et al.* Identical phenotypes of CatSper1 and CatSper2 null sperm. *J. Biol. Chem.* **280**, 32238–32244 (2005).
15. Carlson, A. E. *et al.* CatSper1 required for evoked Ca^{2+} entry and control of flagellar function in sperm. *Proc. Natl Acad. Sci. USA* **100**, 14864–14868 (2003).
16. Qi, H. *et al.* All four CatSper ion channel proteins are required for male fertility and sperm cell hyperactivated motility. *Proc. Natl Acad. Sci. USA* **104**, 1219–1223 (2007).
17. Kaupp, U. B., Kashikar, N. D. & Weyand, I. Mechanisms of sperm chemotaxis. *Annu. Rev. Physiol.* **70**, 93–117 (2008).
18. Xia, J., Reigada, D., Mitchell, C. H. & Ren, D. CATSPER channel-mediated Ca^{2+} entry into mouse sperm triggers a tail-to-head propagation. *Biol. Reprod.* **77**, 551–559 (2007).
19. Olson, S. D., Suarez, S. S. & Fauci, L. J. A model of CatSper channel mediated calcium dynamics in mammalian spermatozoa. *Bull. Math. Biol.* **72**, 1925–1946 (2010).
20. Xia, J. & Ren, D. Egg coat proteins activate calcium entry into mouse sperm via CATSPER channels. *Biol. Reprod.* **80**, 1092–1098 (2009).
21. Huang, L. *et al.* NNC 55–0396 [(1S,2S)-2-(2-(N-[(3-benzimidazol-2-yl)propyl]-N-methylamino)ethyl)-6-fluoro-1,2,3,4-tetrahydro-1-isopropyl-2-naphthyl cyclopropanecarboxylate dihydrochloride]: a new selective inhibitor of T-type calcium channels. *J. Pharmacol. Exp. Ther.* **309**, 193–199 (2004).
22. Baulieu, E. E. Contragestion and other clinical applications of RU 486, an antiprogesterone at the receptor. *Science* **245**, 1351–1357 (1989).
23. Neri-Vidaurri Pdel, C., Torres-Flores, V. & Gonzalez-Martinez, M. T. A remarkable increase in the pH_i sensitivity of voltage-dependent calcium channels occurs in human sperm incubated in capacitating conditions. *Biochem. Biophys. Res. Commun.* **343**, 105–109 (2006).
24. Aitken, R. J., Irvine, S. & Kelly, R. W. Significance of intracellular calcium and cyclic adenosine 3',5'-monophosphate in the mechanisms by which prostaglandins influence human sperm function. *J. Reprod. Fertil.* **77**, 451–462 (1986).
25. Schaefer, M., Hofmann, T., Schultz, G. & Gudermand, T. A new prostaglandin E receptor mediates calcium influx and acrosome reaction in human spermatozoa. *Proc. Natl Acad. Sci. USA* **95**, 3008–3013 (1998).
26. Shimizu, Y. *et al.* Prostaglandins induce calcium influx in human spermatozoa. *Mol. Hum. Reprod.* **4**, 555–561 (1998).
27. Ren, D. & Xia, J. Calcium signaling through CatSper channels in mammalian fertilization. *Physiology (Bethesda)* **25**, 165–175 (2010).
28. Baldi, E. *et al.* Nongenomic activation of spermatozoa by steroid hormones: facts and fictions. *Mol. Cell. Endocrinol.* **308**, 39–46 (2009).
29. Gellersen, B., Fernandes, M. S. & Brosens, J. J. Non-genomic progesterone actions in female reproduction. *Hum. Reprod. Update* **15**, 119–138 (2009).
30. Liu, J., Xia, J., Cho, K. H., Clapham, D. E. & Ren, D. CatSper β , a novel transmembrane protein in the CatSper channel complex. *J. Biol. Chem.* **282**, 18945–18952 (2007).
31. Strünker, T. *et al.* The CatSper channel mediates progesterone-induced Ca^{2+} influx in human sperm. *Nature* doi:10.1038/nature09769 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by the UCSF Program for Breakthrough Biomedical Research.

Author Contributions P.V.L. and Y.K. conceived the project, designed the experiments and wrote the manuscript. P.V.L. performed most of the experiments. I.L.B. helped with pilot experiments for the project. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.K. (yuriy.kirichok@ucsf.edu).

METHODS

Materials and reagents. PGE₁, PGA₁ and PGF_{1 α} were obtained from Avanti Polar Lipids. Oestradiol, PGD₂ and PGE₂ were purchased from Cayman Biochemicals. Progesterone was obtained from Sigma and Calbiochem. Progesterone conjugates were from aBiox. NNC 55-0396 was from Tocris. All other chemicals were purchased from Sigma.

Isolation of spermatozoa. Protocols for the human sperm studies were approved by the Committee on Human Research at the University of California, San Francisco. Freshly ejaculated sperm samples were obtained from four healthy young donors by masturbation and allowed to liquefy for 30–60 min at 22 °C before processing. Human spermatozoa were purified by the swim-up method in artificial human tubal fluid solution (in mM): 98 NaCl, 4.7 KCl, 0.3 KH₂PO₄, 2 CaCl₂, 0.2 MgSO₄, 21 HEPES, 3 glucose, 21 lactic acid, 0.3 sodium pyruvate, pH 7.4 (NaOH). Spermatozoa were stored in this medium at 22 °C for up to 6 h. For *in vitro* capacitation, isolated human spermatozoa were incubated in the capacitating medium as previously described¹³. Electrophysiological properties of non-capacitated and capacitated spermatozoa from the same donor were compared during the same day. Mouse spermatozoa were isolated from cauda epididymis as previously described¹³.

Patch-clamp recordings. Gigaohm seals between the patch pipette and mouse spermatozoa were formed at the cytoplasmic droplet. Human spermatozoa were patched either at the cytoplasmic droplet or, if the cytoplasmic droplet was inconspicuous, at the neck region. Seals were formed in HS solution comprising the following (in mM): 130 NaCl, 5 KCl, 1 MgSO₄, 2 CaCl₂, 5 glucose, 1 sodium pyruvate, 10 lactic acid, 20 HEPES, pH 7.4 adjusted with NaOH. Transition into the whole-cell mode was performed by applying short (1 ms) 499–611 mV voltage pulses, sometimes combined with light suction. Access resistance was 25–70 M Ω , depending on the intracellular solution used (lower for Cs-based compared with NMDG-based solutions). Cells were stimulated every 5 s. Data were sampled at 2–5 kHz and filtered at 1 kHz.

Pipettes (11–17 M Ω) for whole-cell patch-clamp recordings of monovalent CatSper currents were filled with the following (in mM): 130 Cs-methanesulphonate, 70 HEPES/MES, 3 EGTA, 2 EDTA, 0.5 TrisHCl, pH 6.0–8.0 adjusted with CsOH. Bath divalent-free solution for recording of monovalent CatSper currents contained the following (in mM): 140 Cs-methanesulphonate, 40 HEPES/MES, 1 EDTA, pH 7.4 adjusted with CsOH. HS solution was used to record baseline current while measuring monovalent CatSper currents (Ca²⁺ contained in HS solution inhibits monovalent CatSper currents and causes Ca²⁺-dependent inactivation of CatSper channels). Bath solution for recording monovalent CatSper currents at different bath [Ca²⁺]_{free} contained the following (in mM): 140 Cs-methanesulphonate, 40 HEPES/MES, 1 EDTA, 1 BAPTA, 1 HEDTA, pH 7.4 adjusted with CsOH. CaCl₂ was added to this solution in accordance with WinMAXC version 2.05 (C. Patton, Stanford University) to obtain required free [Ca²⁺].

Pipettes (20–30 M Ω) for whole-cell patch-clamp recordings of Ba²⁺ CatSper currents were filled with the following (in mM): 145 NMDG, 100 HEPES, 10 BAPTA, 0.5 TrisHCl, pH 7.4 with methanesulphonic acid (Fig. 3c and Supplementary Fig. 9). Ba²⁺ currents in Supplementary Fig. 2 were obtained with pipette solution (in mM): 140 NMDG, 100 HEPES, 5 EDTA, 5 EGTA, 0.5 TrisHCl, pH 7.0 with methanesulphonic acid. Bath solution for recordings of Ba²⁺ currents contained the following (in mM): 0–50 Ba²⁺, 150–90 NMDG, 100 HEPES, pH 7.4 adjusted with methanesulphonic acid. Mg²⁺ bath solution used as a control (baseline) solution while recording Ba²⁺ currents contained the following (in mM): 2 MgCl₂, 150 NMDG, 100 HEPES, pH 7.4 adjusted with methanesulphonic acid.

Proton currents were recorded in divalent-free solution as described in Lishko *et al.*¹³. Osmolarities of all electrophysiological solutions were approximately 321 and 335 mOsm l⁻¹ for bath and pipette solutions, respectively. All electrophysiology experiments were performed at 24 °C. Data were analysed with Origin 7.0 and Clampfit 9.2. Statistical data were calculated as the mean \pm s.e.m., and *n* indicates number of experiments.

Chromosome length influences replication-induced topological stress

Andreas Kegel¹, Hanna Betts-Lindroos¹, Takaharu Kanno¹, Kristian Jeppsson¹, Lena Ström¹, Yuki Katou², Takehiko Itoh², Katsuhiko Shirahige³ & Camilla Sjögren¹

During chromosome duplication the parental DNA molecule becomes overwound, or positively supercoiled, in the region ahead of the advancing replication fork. To allow fork progression, this superhelical tension has to be removed by topoisomerases, which operate by introducing transient DNA breaks¹. Positive supercoiling can also be diminished if the advancing fork rotates along the DNA helix, but then sister chromatid intertwinings form in its wake^{1,2}. Despite these insights it remains largely unknown how replication-induced superhelical stress is dealt with on linear, eukaryotic chromosomes. Here we show that this stress increases with the length of *Saccharomyces cerevisiae* chromosomes. This highlights the possibility that superhelical tension is handled on a chromosome scale and not only within topologically closed chromosomal domains as the current view predicts. We found that inhibition of type I topoisomerases leads to a late replication delay of longer, but not shorter, chromosomes. This phenotype is also displayed by cells expressing mutated versions of the cohesin- and condensin-related Smc5/6 complex. The frequency of chromosomal association sites of the Smc5/6 complex increases in response to chromosome lengthening, chromosome circularization, or inactivation of topoisomerase 2, all having the potential to increase the number of sister chromatid intertwinings³. Furthermore, non-functional Smc6 reduces the accumulation of intertwinings sister plasmids after one round of replication in the absence of topoisomerase 2 function. Our results demonstrate that the length of a chromosome influences the need of superhelical tension release in *Saccharomyces cerevisiae*, and allow us to propose a model where the Smc5/6 complex facilitates fork rotation by sequestering nascent chromatid intertwinings that form behind the replication machinery.

The Smc5/6 complex (hereafter referred to as Smc5/6) consists of Smc5, Smc6, Nse1, the Sumo-ligase Mms21 and Nse3–Nse6 (ref. 4), and has to be functional during replication for normal mitotic progression⁵. The exact role of this complex in unchallenged cells is largely unknown. Notably, chromatin immunoprecipitation (ChIP) on chip analysis of cells arrested at G2/M shows that the frequency of arm-binding sites increases with chromosome length in budding yeast⁵. Using ChIP sequencing, which allows a more quantitative determination of a protein-binding profile⁶, we confirmed the length dependency of Smc5/6 chromosome association (Supplementary Fig. 2 and Supplementary Table 2). As this binding pattern shows no consistent correlation with established chromosomal features, we hypothesized that it was a reflection of replication-induced superhelical stress, and that this increases with chromosome length. If so, topoisomerases that remove this tension are expected to have a length-dependent effect on replication. Type I topoisomerases (Top1 and Top3) release superhelical tension by introducing temporary nicks in the DNA, whereas type II topoisomerases (Top2) transfer one DNA double helix through another via transient formation of a double-strand break (DSB)¹.

Either Top1 or Top2 can support replication in budding yeast, but Top1 seems to have the central function in the release of supercoils ahead of the fork^{7,8}. Top2, on the other hand, is central for removal of sister chromatid intertwinings (SCIs), which is a prerequisite for correct chromosome segregation (Fig. 1a)^{3,7,8}. The function of Top3 on undamaged chromosomes is unclear, but it has been suggested to remove the topological strain between two converging replication forks^{1,9,10}. The growth of *top3Δ* mutants is severely hampered, probably due to the accumulation of aberrant chromosome structures

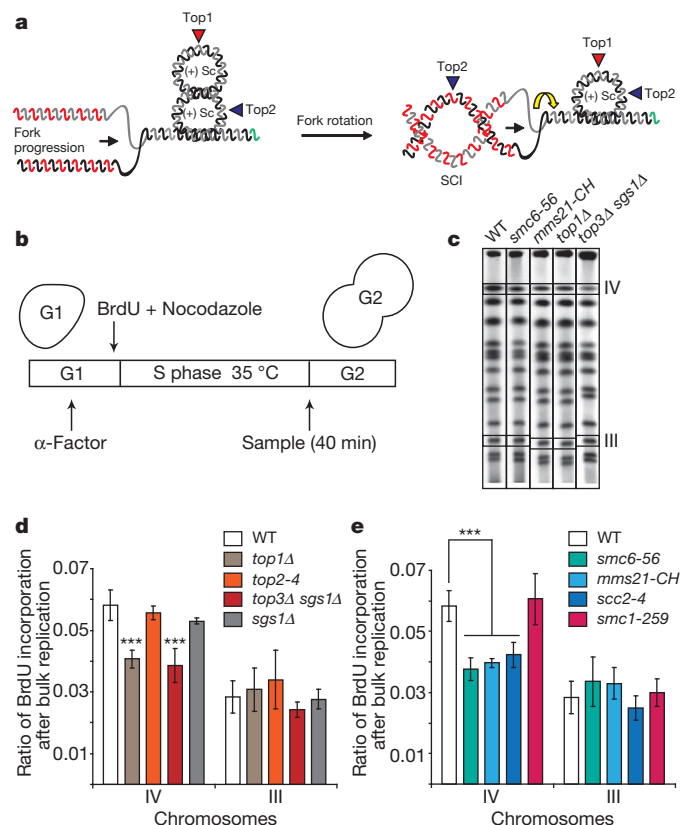


Figure 1 | Top1, Top3 and Smc5/6 are required for timely completion of replication on long chromosomes. **a**, Release of replication-induced superhelical tension by topoisomerases or fork rotation. Arrow, replication direction; red/blue arrowheads, topoisomerase 1/2; yellow arrow, fork rotation; (+) Sc, positive supercoil; SCI, sister chromatid intertwinings. See also Supplementary Fig. 1. **b**, Illustration of chromosomal BrdU labelling during S phase. **c**, Immunodetection of BrdU incorporation after chromosome separation by PFGE. Chromosomes IV and III are highlighted. **d**, **e**, Quantification of BrdU-labelled chromosomes. The signals for Chr IV and III were normalized to total BrdU incorporation. Standard deviations and *P*-values (*t*-test, ****P* < 10⁻³) are based on *n* = 27 (wild type) and *n* = 3 (mutants).

¹Karolinska Institutet, Department of Cell and Molecular Biology, von Eulers väg 3, 171 77 Stockholm, Sweden. ²Laboratory of In Silico Functional Genomics, Graduate School of Bioscience, Tokyo Institute of Technology, 4259 Nagatsuta, Midoriku, Yokohama 226-8501, Japan. ³Research Center for Epigenetic Disease, Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan.

created by the Sgs1 helicase¹¹. Top3 function is therefore commonly addressed in *top3Δ sgs1Δ* cells, which grow at wild-type rates.

To test if any of the topoisomerases has a length-dependent function, chromosomes from wild-type, *top1Δ*, *top3Δ sgs1Δ*, *sgs1Δ* and temperature-sensitive *top2-4* cells were isolated after one round of replication at 35 °C in the presence of 5-bromodeoxyuridine (BrdU) (Fig. 1 and Supplementary Fig. 3a). After separation by pulse-field gel electrophoresis (PFGE), blotting and detection using anti-BrdU antibodies (Fig. 1c), the signals for a short (Chr III, ~317 kb) and a long (Chr IV, ~1,532 kb) chromosome were quantified and compared to the signal of total BrdU incorporation. Although the relative amount of Chr III that migrated into the gel was similar in wild-type and mutated cells, significantly less of Chr IV was detected in *top1Δ* and *top3Δ sgs1Δ* cells, but not in *sgs1Δ* or *top2-4* cells (Fig. 1d). Such delay was detected for all long chromosomes down to a size of 924 kb (Chr XIII) (Supplementary Fig. 4). Because inhibition of chromosome penetration into pulse-field gels depends on the branched structures of ongoing replication forks¹², this shows that Top1 and Top3 are needed for the timely completion of replication of longer chromosomes. As the fluorescence-activated cell sorting (FACS) profiles show that bulk replication is largely unperturbed, this delay probably occurs late in S phase (Supplementary Fig. 3a). This is in line with a function of Top3 in replication termination, and also indicates that Top1 is important at later stages of replication^{1,9,10}.

In further support of a length-dependent function of Top1, we confirmed that inhibition of Top1 by camptothecin (CPT) influences replication of longer yeast chromosomes in particular¹³ (Supplementary Fig. 5). CPT inhibition of Top1 is generally thought to interfere with the DNA re-ligation step¹⁴, causing replication fork blocks and checkpoint activation. However, CPT also inhibits DNA nicking by Top1 (ref. 15), and increases supercoiling ahead of a replication fork¹⁶. Furthermore, the yeast damage checkpoint remains inactivated when CPT is present during a single S phase as in the experiment performed here¹³. This indicates that the length-dependent replication delay in CPT-treated cells is at least in part caused by unresolved topological tension in the DNA molecule. Together, these results indicate that topological stress increases with chromosome length.

The absence of a replication delay in *top2-4* mutants confirms earlier reports showing that unresolved SCIs leave replication largely unperturbed^{7,8}, and demonstrates that SCIs are not the cause of the PFGE phenotype detected here.

Having established that superhelical tension accumulates on longer chromosomes, we analysed temperature-sensitive *smc6-56* and SUMO ligase-dead *mms21-CH* mutants in the PFGE/BrdU assay. Both mutations triggered a length-dependent late replication delay, indicating that Smc5/6 also handles replication-induced topological stress (Fig. 1e). Scc2, which is needed for loading of both cohesin and Smc5/6 to undamaged chromosomes^{5,17}, was also required for full replication of long chromosomes (Fig. 1e). In contrast, replication was unperturbed in the cohesin mutant *smc1-259*, indicating that the effect on replication is Smc5/6-specific (Fig. 1e). In *smc6-56* cells, replication was significantly inhibited on chromosomes which were ~746 kb or longer (Supplementary Fig. 4). The reason why the *smc6-56* replication defect extends to chromosomes of shorter length than *top1Δ* and *top3Δ sgs1Δ* is unclear. Possibly, the absence of Smc5/6 function triggers a higher accumulation of topological tension than removal of the topoisomerases, and/or the lack of topoisomerases is more easily compensated.

In addition to its function in undamaged cells, the Smc5/6 complex prevents accumulation of aberrant recombination structures at stalled and collapsed replication forks^{18–20}. To test whether the replication delay detected in the *smc6-56* mutant was due to faulty recombination, the PFGE/BrdU assay was repeated using cells lacking the central recombination proteins Rad51 or Rad52. The *smc6-56* phenotype was unchanged in *smc6-56 rad51Δ* and *smc6-56 rad52Δ* cells, showing that the length-dependent function is unrelated to aberrant recombination (Supplementary Fig. 6a, b, data not shown). Furthermore, chromosome segregation was equally delayed in *smc6-56* and *smc6-56 rad51Δ* cells (Supplementary Fig. 7a–d). Because earlier observations show that Smc5/6 is needed during replication for correct segregation, this provides yet another argument that the complex's S-phase function is unrelated to aberrant recombination. Additionally, cells lacking Srs2, which negatively regulates recombination^{21,22}, or Sgs1, which together with Top3 and Rmi1 is involved in the resolution of recombination structures^{23,24}, did not display a chromosome-length-dependent

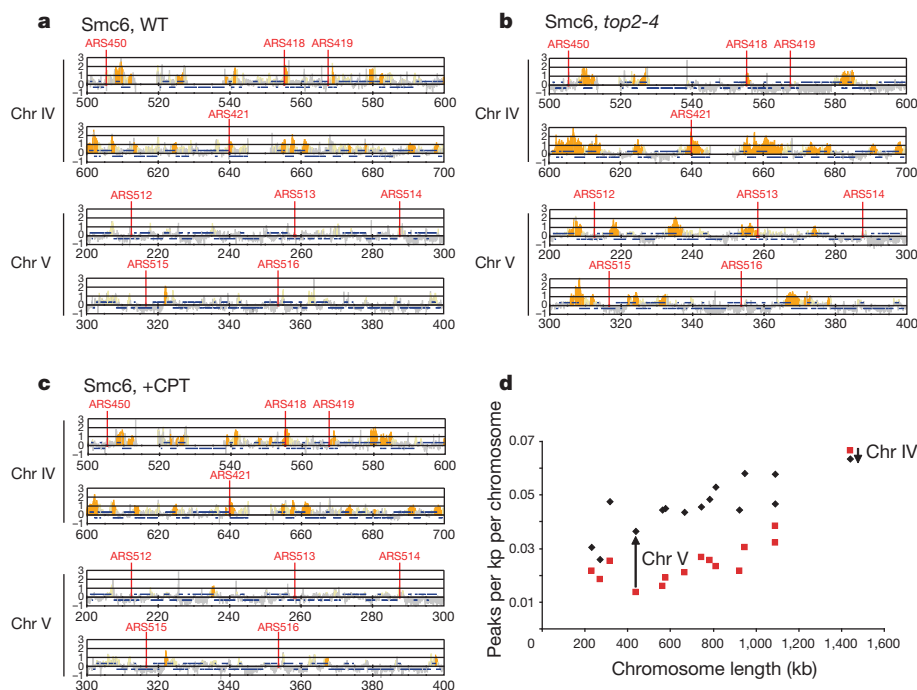


Figure 2 | Inactivation of Top2, but not Top1, increases the frequency of Smc6 chromosomal interactions. a–c, Chromosomal association of Flag-tagged Smc6 in wild-type (a), *top2-4* (b) and wild-type CPT-treated (c) cells. Orange peaks, significant chromosomal binding sites; blue horizontal lines,

open reading frames. The y axis shows log₂ of signal strength; the x axis shows chromosomal coordinates in kilobases. Vertical red lines indicate replication origins. d, Correlation between the number of Smc6 binding sites per kb and chromosome length. Red squares, wild type; black diamonds, *top2-4*.

phenotype in the PFGE assay (Fig. 1d and Supplementary Fig. 6c, d). Taken together, this argues that the length-dependent replication delay, and the defect in chromosome segregation, is independent of aberrant recombinational repair of spontaneous DNA breaks.

The similar replication delay found in type 1 topoisomerase mutants and Smc5/6 mutants is in line with a role for the complex in resolving replication-induced topological tension. To test if the chromosomal association of Smc5/6 is a reflection of this stress, we analysed whether it was dependent on replication as such, and not merely passage through S phase. Cell cycle progression without initiation of replication was achieved by depletion of Cdc6, a central component of the pre-replication complex²⁵. Under these conditions, the number of Smc6 chromosomal interaction sites in G2/M cells decreased by ~60%, showing that the association is determined by the replication process (Supplementary Fig. 8).

To test further the assumption that the chromosomal association of Smc5/6 is determined by chromosome topology, Smc6 chromosomal association was analysed after Top1 or Top2 inactivation. CPT inhibition of Top1 did not significantly alter the Smc6 interaction pattern, but an S phase at non-permissive temperature for *top2-4* triggered a marked change (Fig. 2a–c). The overall number of Smc6 binding sites increased by 92%. Because inhibition of *top2-4* leaves replication unperturbed but increases the number of SCIs⁷, this indicates that

Smc5/6 association is triggered by these. Interestingly, the biggest increase of Smc6 interaction sites was detected on chromosomes of intermediate length, and even though the Smc6 signals on the long Chr IV seem stronger in *top2-4* cells, there was little change in the number of association sites (Fig. 2a, b, d). We interpret this as follows. As proposed previously, SCIs are able to move along chromosomes, and this movement is easier if the chromosome is short³. Owing to a lower mobility of SCIs on longer chromosomes, they accumulate and trigger Smc5/6 interaction even in wild-type cells with functional Top2. On chromosomes of intermediate length, SCIs are less stable in wild-type cells but accumulate and lead to more Smc5/6 interaction sites after Top2 inhibition. Finally, on the shortest chromosomes, the SCIs are free to swivel off chromosome ends, and few Smc5/6-binding sites are detected in both wild-type and *top2-4* cells. The absence of an increase in Smc6 binding on Chr IV in *top2-4* cells could indicate that the frequency of SCIs has reached a level of saturation. Interestingly, the ChIP-seencing data also indicate that the amount of Smc6 levels off when reaching this chromosome length (Supplementary Fig. 2a).

If the above is true, Smc5/6 is expected to disappear from a long chromosome that is shortened, and to accumulate on a circular variant of a short chromosome. Indeed, in cells where the ~1,532-kb-long Chr IV was present as two pieces of ~500 and ~1,032 kb, the frequency of Smc6 binding sites decreases to the levels detected on natural

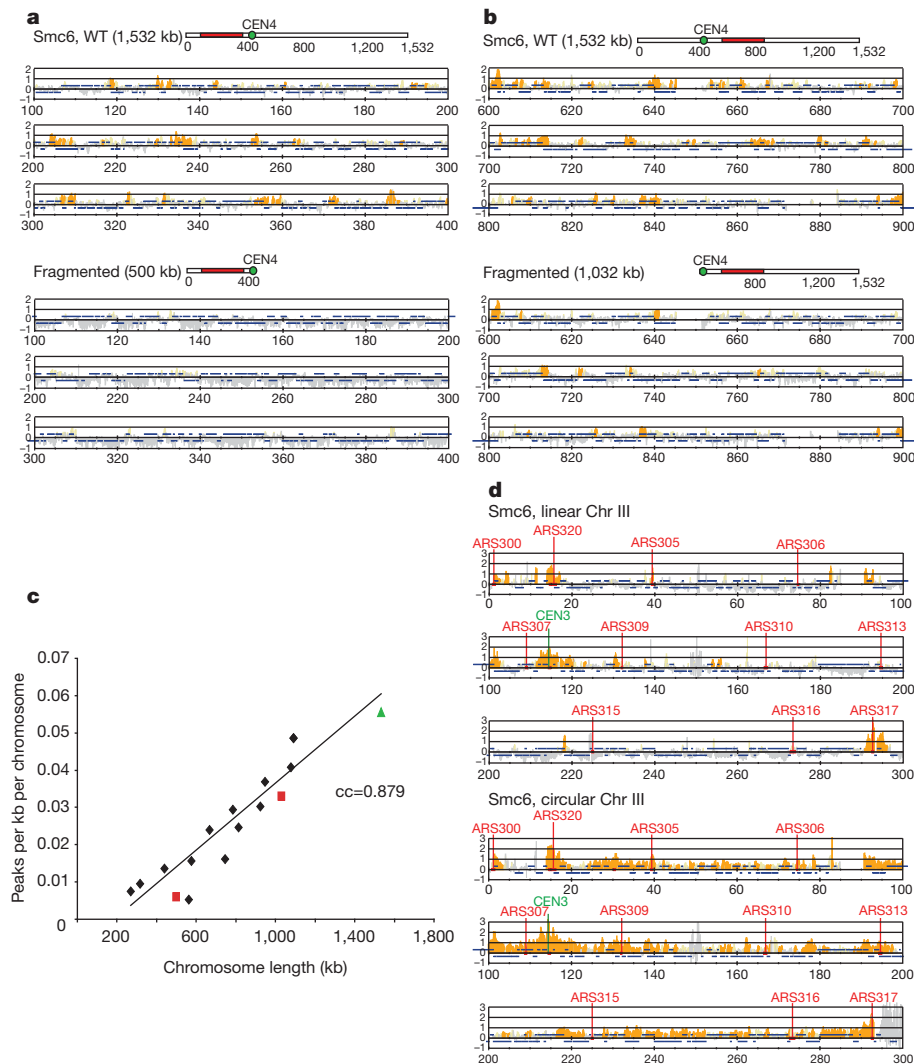


Figure 3 | The Smc5/6 complex senses chromosome length and circularization of Chr III. a, b, d, Chromosomal association of Flag-tagged Smc6. Annotation as in Fig. 2. a, b, Normal or fragmented Chr IV; the area highlighted in red shows the chromosomal region displayed in the localization

map. d, Linear or circular Chr III. c, Correlation between the number of Smc6 binding sites per kb and chromosome length. Black diamonds, Chr II–III and Chr V–XVI; green triangle, normal Chr IV (~1,532 kb); red squares, fragmented Chr IV (~500 and ~1,032 kb); cc, correlation coefficient.

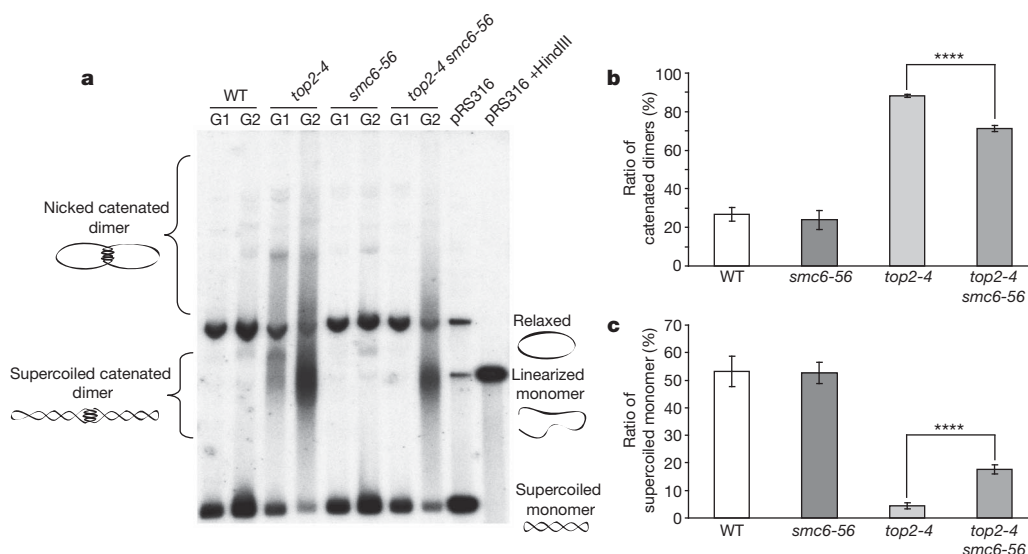


Figure 4 | The Smc5/6 complex facilitates catenation of an episomal plasmid. **a**, Southern blot analysis of pRS316 from indicated cells. Controls for relaxed, supercoiled and linear monomers are shown in the two rightmost lanes. Nicked and supercoiled catenated dimers, as well as supercoiled

chromosomes of the same size (Fig. 3a–c)²⁶. In addition, the number of both Smc6 and Nse4 interaction sites increased markedly on a circular variant of Chr III (Chr III^{circ}) (Fig. 3d and Supplementary Fig. 9a). ChIP sequencing confirmed that the increase in binding frequency reflected a higher amount of Smc6 on Chr III^{circ} (Supplementary Fig. 2a–c). The accumulation was not due to an increased number of DNA breaks, as Smc6 still amasses on this chromosome in cells lacking Mre11, which is needed for the localization of the complex to DSBs⁵ (Supplementary Fig. 9b). The changed distribution of Smc6 in *top2-4* cells, and after chromosome shortening or circularization, argues that Smc5/6 loading is triggered by SCIs.

To gain further insight into how the complex is recruited to chromosomes, the interaction pattern of the loading protein Scc2 was analysed. Scc2 shows few additive interaction sites on Chr III^{circ}, and it does not display a length-dependent interaction pattern (Supplementary Figs 9e and 10). This indicates that Smc5/6 is able to move away from its loading sites like cohesin²⁷, and might finally accumulate at SCIs. Neither Smc1 nor Smc2—subunits of cohesin and condensin complexes, respectively—change their interaction pattern in response to Chr III circularization (Supplementary Fig. 9c, d). This indicates that among eukaryotic SMC complexes only Smc5/6 associates to sites of SCIs.

How can a chromosomal localization process dependent on SCI formation be reconciled with a type-I-topoisomerase-like function for Smc5/6? Possibly, the complex binds and sequesters nascent SCIs behind the replication fork. This could drive fork rotation and thereby the release of superhelical tension ahead of the replication machinery (Supplementary Fig. 1). If true, non-functional Smc5/6 is expected to trigger an increase in supercoiling and decrease the number of SCIs. To test this, the topological status of a 4.9-kb plasmid was investigated using gel electrophoresis and Southern blot analysis^{28,29}. The plasmid was introduced into wild-type, *smc6-56*, *top2-4*, or *top2-4 smc6-56* cells, and isolated from cells arrested in G1 at 25 °C or in G2 after a single passage through S phase at non-permissive temperature for the *smc6-56* and *top2-4* alleles. Without the use of *top2-4*, the effect of *smc6-56* on SCI formation is not possible to study, as very few SCIs accumulate in wild-type cells. As shown before, SCIs accumulated in the G2 sample of *top2-4* cells, increasing the amount of catenated plasmid dimers²⁸ (Fig. 4a, b and Supplementary Figs 3b and 11). This increase was significantly diminished in *top2-4 smc6-56* double mutants, and in line with a defect in transferring supercoils to SCIs, the amount of supercoiled, monomeric plasmids was increased (Fig. 4a–c).

monomers, are indicated (see Supplementary Fig. 11 and ref. 28).

b, c, Percentage of catenated dimers (**b**) and supercoiled monomers (**c**) in the G2/M-arrested samples. Standard deviations and *P* values (*t*-test, *****P* < 10^{−5}) are based on results from *n* = 4 experiments.

Supercoiled material did not accumulate in the single *smc6-56* mutant, probably due to a compensatory function of the topoisomerases still present in these cells (Fig. 4c).

The presented data favour a model where Smc5/6 reduces the accumulation of positive supercoils ahead of the replication machinery by assisting fork rotation through sequestration of SCIs (Supplementary Fig. 1). Possibly the complex, via its association to SCIs, also organizes the chromatids in a ‘rotation friendly’ conformation. The length-dependent role of Smc5/6 and type I topoisomerases could be explained if the rotation occurs on a chromosome scale, and thus is easier to accomplish when a chromosome is short (Supplementary Fig. 1). This would increase the efficacy of Smc5/6 on short chromosomes, leaving fewer supercoils for the topoisomerases to handle. Alternatively, detectable replication delays occur with a fixed probability of less than one per replicon, and as there are more replicons on a long chromosome, the risk increases for a discernible fork blockage. Future studies testing the topological tension in linear chromosomes of different length will determine this.

Although the rotation model unites the findings presented here, other modes of function for Smc5/6 are still possible. The complex could, for example, activate type I topoisomerases directly, and its binding pattern and connection to SCI formation and/or stabilization could be due to more passive roles of the complex in organizing chromosome structure, as suggested previously³⁰. Regardless, the Smc5/6 complex, via its interplay with topology and topoisomerases, must have a key role in the structural organization of chromosomes. Furthermore, if it channels superhelical tension ahead of the fork into SCIs behind, as our results indicate, the strength of superhelical tension can be a key determinant in the organization of chromosome structure.

METHODS SUMMARY

The yeast strains used in this work are haploid and derivatives of W303, YPH952 (ref. 26), or R071 (see Supplementary Table 1 for details). Methods used for ChIP-on-chip, ChIP sequencing, PFGE/BrdU, and plasmid topology analysis are described in the Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 November 2010; accepted 5 January 2011.

Published online 2 March 2011.

- Wang, J. C. Cellular roles of DNA topoisomerases: a molecular perspective. *Nature Rev. Mol. Cell Biol.* **3**, 430–440 (2002).

2. Postow, L., Crisona, N. J., Peter, B. J., Hardy, C. D. & Cozzarelli, N. R. Topological challenges to DNA replication: conformations at the fork. *Proc. Natl Acad. Sci. USA* **98**, 8219–8226 (2001).
3. Spell, R. M. & Holm, C. Nature and distribution of chromosomal intertwinings in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **14**, 1465–1476 (1994).
4. Hazbun, T. R. et al. Assigning function to yeast proteins by integration of technologies. *Mol. Cell* **12**, 1353–1365 (2003).
5. Lindroos, H. B. et al. Chromosomal association of the Smc5/6 complex reveals that it functions in differently regulated pathways. *Mol. Cell* **22**, 755–767 (2006).
6. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **316**, 1497–1502 (2007).
7. Bermejo, R. et al. Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes Dev.* **21**, 1921–1936 (2007).
8. Kim, R. A. & Wang, J. C. Function of DNA topoisomerases as replication swivels in *Saccharomyces cerevisiae*. *J. Mol. Biol.* **208**, 257–267 (1989).
9. Hiasa, H., DiGate, R. J. & Marians, K. J. Decatenating activity of *Escherichia coli* DNA gyrase and topoisomerases I and III during oriC and pBR322 DNA replication *in vitro*. *J. Biol. Chem.* **269**, 2093–2099 (1994).
10. Mankouri, H. W. & Hickson, I. D. The RecQ helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific 'dissolvosome'? *Trends Biochem. Sci.* **32**, 538–546 (2007).
11. Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L. & Rothstein, R. The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol. Cell. Biol.* **14**, 8391–8398 (1994).
12. Hennessy, K. M., Lee, A., Chen, E. & Botstein, D. A group of interacting yeast DNA replication genes. *Genes Dev.* **5**, 958–969 (1991).
13. Redon, C. et al. Yeast histone 2A serine 129 is essential for the efficient repair of checkpoint-blind DNA damage. *EMBO Rep.* **4**, 678–684 (2003).
14. Pommier, Y. et al. Repair of topoisomerase I-mediated DNA damage. *Prog. Nucleic Acid Res. Mol. Biol.* **81**, 179–229 (2006).
15. Kjeldsen, E., Svejstrup, J. Q., Gromova, I. I., Alsner, J. & Westergaard, O. Camptothecin inhibits both the cleavage and religation reactions of eukaryotic DNA topoisomerase I. *J. Mol. Biol.* **228**, 1025–1030 (1992).
16. Koster, D. A., Palle, K., Bot, E. S., Bjornst, M. A. & Dekker, N. H. Antitumour drugs impede DNA uncoiling by topoisomerase I. *Nature* **448**, 213–217 (2007).
17. Ciosk, R. et al. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol. Cell* **5**, 243–254 (2000).
18. Ampatzidou, E., Irmisch, A., O'Connell, M. J. & Murray, J. M. Smc5/6 is required for repair at collapsed replication forks. *Mol. Cell. Biol.* **26**, 9387–9401 (2006).
19. Branzei, D. et al. Ubc9- and mms21-mediated sumoylation counteracts recombinogenic events at damaged replication forks. *Cell* **127**, 509–522 (2006).
20. Sollier, J. et al. The *Saccharomyces cerevisiae* Esc2 and Smc5–6 proteins promote sister chromatid junction-mediated intra-S repair. *Mol. Biol. Cell* **20**, 1671–1682 (2009).
21. Krejci, L. et al. DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* **423**, 305–309 (2003).
22. Veaute, X. et al. The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. *Nature* **423**, 309–312 (2003).
23. Gangloff, S., Soustelle, C. & Fabre, F. Homologous recombination is responsible for cell death in the absence of the Sgs1 and Srs2 helicases. *Nature Genet.* **25**, 192–194 (2000).
24. Liberi, G. et al. Rad51-dependent DNA structures accumulate at damaged replication forks in *sgs1* mutants defective in the yeast ortholog of BLM RecQ helicase. *Genes Dev.* **19**, 339–350 (2005).
25. Cocker, J. H., Piatti, S., Santocanale, C., Nasmyth, K. & Diffley, J. F. An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature* **379**, 180–182 (1996).
26. Hamer, L., Johnston, M. & Green, E. D. Isolation of yeast artificial chromosomes free of endogenous yeast chromosomes: Construction of alternate hosts with defined karyotypic alterations. *Proc. Natl Acad. Sci. USA* **92**, 11706–11710 (1995).
27. Lengronne, A. et al. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**, 573–578 (2004).
28. Baxter, J. & Diffley, J. F. Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast. *Mol. Cell* **30**, 790–802 (2008).
29. Koshland, D. & Hartwell, L. H. The structure of sister minichromosome DNA before anaphase in *Saccharomyces cerevisiae*. *Science* **238**, 1713–1716 (1987).
30. Outwin, E. A., Irmisch, A., Murray, J. M. & O'Connell, M. J. Smc5-Smc6-dependent removal of cohesin from mitotic chromosomes. *Mol. Cell. Biol.* **29**, 4363–4375 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Nasmyth, J. Haber, E. Green and X. Zhao for yeast strains and the BEA core facility at Karolinska Institutet for help with ChIP on chip. Financial support: Strategic Japanese-Swedish Cooperative Program from JST, SSF and Vinnova (C.S. and K.S.); please see Supplementary Information for additional support.

Author Contributions A.K. performed the PFGE-based assays; A.K. and T.K. the plasmid assays; A.K., H.B.-L. and K.J. the ChIP-on-chip; K.J., Y.K. and K.S. the ChIP sequencing. T.I. and K.S. carried out the computational analysis and C.S. the segregation experiment. H.B.-L., L.S., K.S. and C.S. initiated the study, A.K., K.S. and C.S. continued and finalized its design. A.K. and C.S. wrote the paper. All authors analysed data, discussed the results and commented on the manuscript.

Author Information Original data files from ChIP-sequencing experiments can be found at <http://trace.ncbi.nlm.nih.gov/Traces/sra/>, accession number SRP004920, and from ChIP-on-chip experiments at <http://www.ncbi.nlm.nih.gov/geo/>, accession number GSE26263. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.S. (camilla.sjogren@ki.se).

METHODS

The yeast strains used in this work are haploid and derivatives of W303, YPH952 (ref. 26), or R071 (see Supplementary Table 1 for details). Cell cycle arrests, FACS and Southern blot analyses were performed using established techniques.

ChIP-on-chip and ChIP sequencing. ChIP-on-chip was performed as described previously^{5,31} using *S. cerevisiae* whole-genome tiling 1.0F arrays (Affymetrix). For ChIP-sequencing analysis, more than 10 million tag sequences for both ChIP and whole-cell extract (WCE) fractions were read. Sequence coverage is higher than 92% and total redundancy is over 30-fold in all cases (Supplementary Table 2). Chromatin immunoprecipitation was carried out as previously described³². DNA from the WCE and ChIP fractions was further sheared to the average size of 150 bp by ultra sonicator Covaris (Covaris Inc.), end-repaired, ligated to sequencing adapters and amplified according to the manufacturer's instructions (Applied Biosystems SOLiD Library Preparation Protocol). Gel-purified amplified DNA between 100 and 150 bp was sequenced on the Applied Biosystems Solid 3.5 platform to generate 50-bp reads. Sequence reads were aligned to the *S. cerevisiae* reference genome using Corona_lite (Applied Biosystems) allowing three colour space mismatches. More than 10 million reads were mapped for each sample. Aligned reads were extended to 100 bp in the 3' direction. As inclusion of the repetitive data in ChIP-sequencing experiments is critical for a complete understanding of protein localization along the chromosome³³, sequence reads mapped to repetitive sequences were divided equally among all locations where the repetitive sequence appeared, treating each repetitive sequence as being equally responsible for the resulting sequencing read. The number of reads was summed up in a 500-bp window with a step size of 50 bp along the chromosome for ChIP and WCE fractions, respectively. After normalization of total reads of ChIP fraction against WCE fraction, enrichment values (ChIP/WCE) were calculated as previously described for each window and presented³².

All samples for ChIP-on-chip and ChIP sequencing were collected from benomyl- or nocodazole-arrested ($80 \mu\text{g ml}^{-1}$ and $15 \mu\text{g ml}^{-1}$, respectively) G2/M cells.

Cdc6 depletion. Cdc6 depletion was mainly performed as described previously³⁴. Cells with Flag-tagged Smc6 and the *CDC6* gene under the control of a galactose-inducible promoter grown in medium containing 2% galactose and raffinose were first arrested in G1 using α -factor (Innovagen). Twenty minutes after release from this arrest, half of the culture was washed and re-suspended in medium with 2% raffinose to deplete cells of Cdc6. α -Factor was also added back to both cultures to re-arrest them in subsequent G1. The two cultures were finally released into a nocodazole G2/M arrest, maintaining the presence of raffinose in one, and raffinose and galactose in the other. Samples for PFGE and Southern blot analysis using probes against Chr III and Chr IV were collected every 20 min after the last G1

release. Whole-genome ChIP-on-chip analysis of Smc6 was performed on the G2/M arrested cells, and FACS analysis done every 20 min during the entire experiment. **PFGE/BrdU analysis.** Analysis of replication by PFGE of BrdU-containing chromosomes was performed as described previously³⁵. Cells were grown at 21 °C and arrested in G1 using α -factor. The temperature was raised to 35 °C for an additional 45 min under maintained arrest, and cells were subsequently released into medium containing BrdU ($400 \mu\text{g ml}^{-1}$; Sigma), and nocodazole ($15 \mu\text{g ml}^{-1}$; Sigma) for arrest in G2/M. Samples were collected and analysed as indicated in the text when all cells had reached G2/M after 40 min at 35 °C. CPT ($20 \mu\text{M}$) was added to the G1-arrested cells 30 min before release, and kept in CPT-containing medium during S-phase progression. Preparation of chromosomes for PFGE analysis was performed as described previously⁵.

Plasmid assay. For analysis of the topological variants of a plasmid, strains with pRS316 were arrested in G1 and released into a G2/M arrest at 35 °C as during the PFGE assay and as described in ref. 28. To avoid alterations of plasmid topology owing to sample preparation, cell lysis was performed in agarose plugs as in ref. 36. Southern blot using a radioactive probe against plasmid-specific sequences was performed using standard methods. Gel electrophoresis of plasmid samples was performed on a 0.8% agarose gel in the presence of $0.5 \mu\text{g ml}^{-1}$ ethidium bromide in all cases except in Supplementary Fig. 11a where ethidium bromide was omitted. Ratios of supercoiled or catenated plasmids were calculated using the formula: ratio of supercoiled monomers or catenated dimers = supercoiled monomers or catenated dimers/(supercoiled monomers + relaxed monomers + catenated dimers) $\times 100$. Statistical analyses to generate standard deviations and *P* values were performed using the statistical analysis tools provided by Microsoft Office Excel 2003. The statistical significance of the data represented by the *P* value was determined by *t*-test.

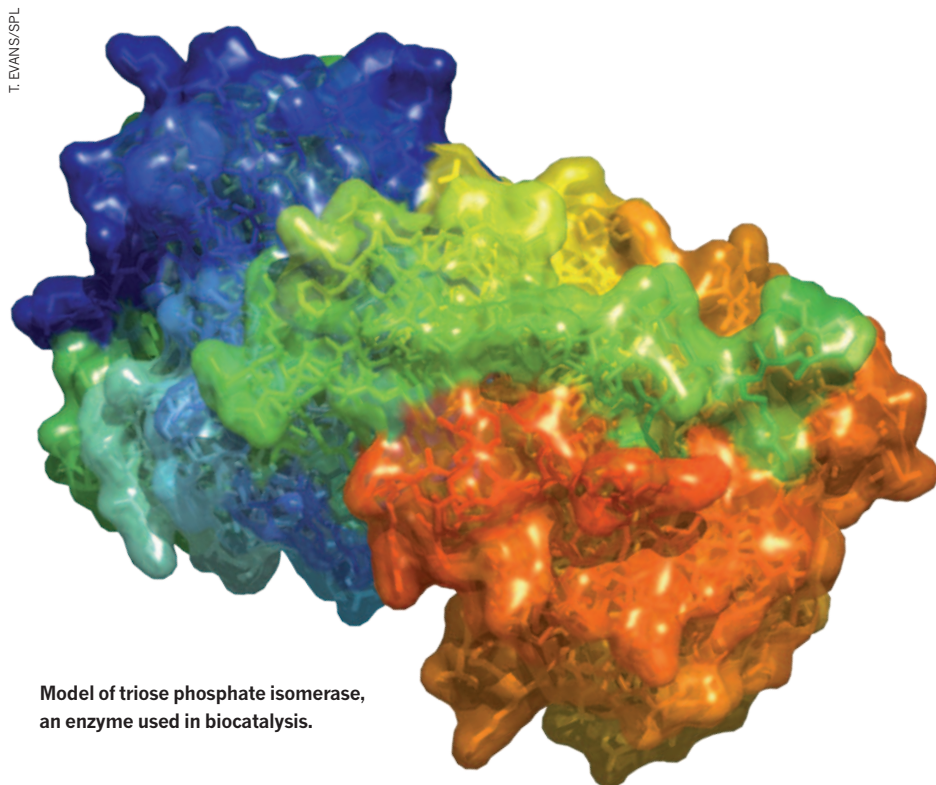
31. Katou, Y., Kaneshiro, K., Aburatani, H. & Shirahige, K. Genomic approach for the understanding of dynamic aspect of chromosome behavior. *Methods Enzymol.* **409**, 389–410 (2006).
32. Katou, Y. *et al.* S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. *Nature* **424**, 1078–1083 (2003).
33. Rosenfeld, S. Characteristics of transcriptional activity in nonlinear dynamics of genetic regulatory networks. *Gene Regul. Syst. Biol.* **3**, 159–179 (2009).
34. Piatti, S., Böhm, T., Cocker, J. H., Diffley, J. F. & Nasmyth, K. Activation of S-phase-promoting CDKs in late G1 defines a 'point of no return' after which Cdc6 synthesis cannot promote DNA replication in yeast. *Genes Dev.* **10**, 1516–1531 (1996).
35. Lengronne, A., Pasero, P., Bensimon, A. & Schwob, E. Monitoring S phase progression globally and locally using BrdU incorporation in TK⁺ yeast strains. *Nucleic Acids Res.* **29**, 1433–1442 (2001).
36. Sjögren, C. & Nasmyth, K. Sister chromatid cohesion is required for postreplicative double-strand break repair in *Saccharomyces cerevisiae*. *Curr. Biol.* **11**, 991–995 (2001).

CAREERS

UNITED STATES NIH granting formula frustrates some senior scientists **p.399**

READER POLL How many hours do postdocs work per week? go.nature.com/6woflx

NATUREJOBS For the latest career listings and advice www.naturejobs.com



Model of triose phosphate isomerase, an enzyme used in biocatalysis.

CHEMISTRY

Enzyme expertise

Biocatalysis specialists are in high demand in industry.

The global appetite for goods, fuel and pharmaceuticals that are derived from petrochemicals is squeezing the planet dry of resources. Chemists may be able to offer solutions by synthesizing products from alternative raw materials. But one of the most promising avenues in the quest for sustainability may be an amalgamation of chemistry and biology known as biocatalysis, in which biologically produced enzymes are used instead of heavy-metal catalysts to improve the efficiency and sustainability of reactions. Enzymes are gaining popularity in industry, and experts in biocatalysis are becoming hot commodities.

Catalysts are required for chemical transformations ranging from turning a crude-oil mixture into chemicals for manufacturing to making a molecular tweak to a drug candidate. Catalysts are traditionally made from metals

such as platinum, rhodium or palladium, with various molecular add-ons designed to help each step of a synthesis. But enzymes can hasten complicated transformations in a single step. By tinkering with enzymes, researchers can coax previously resource-intensive reactions to proceed faster at lower temperatures and pressures, with fewer work-up steps and a higher yield — all of which improves efficiency and saves energy.

Although enzymes cannot replace every catalyst used in industry, they could supplant a great many if companies embrace the technology. Right now, firms seem willing to do this. In the pharmaceutical industry, blighted by thousands of job losses in the past few years, biocatalysis provides a glimmer of hope for scientists seeking opportunities. For example, Novozymes, an enzyme-development

company based in Bagsværd, Denmark, has increased its team of scientists from 581 in 2001 to 961 in 2010. And a small biocatalysis group at the research labs of pharmaceuticals giant Merck in Rahway, New Jersey, has survived three rounds of cuts in the past decade, and looks set to remain core to the company's business. Management, says Jeffery Moore, part of that team, "views biocatalysis as a distinguishing feature of chemistry at Merck and a competitive advantage in the marketplace". He adds that because of the lay-offs, most of his recruits are newly redundant chemists, sometimes from Merck, rather than industry new-comers. Competition is tough.

The group at Merck is small — just seven people — but the company also creates opportunities beyond its own confines. "Pharmaceutical companies are outsourcing a lot of synthetic work to lower-cost manufacturers," says Moore. "Where you see a lot of growth in biocatalysis is at those third-party manufacturers."

CATALYSING OPPORTUNITIES

Biocatalysis requires the skills of chemists who can devise synthetic routes to a given molecule; biochemists who can understand enzyme function and reaction conditions; and microbiologists who can engineer the enzymes. There are also increasing opportunities for chemical engineers with an enzymatic bent, says David Rozzell, a consultant based in Burbank, California, who sold his own biocatalysis company, BioCatalytics, in 2007. "I think that the jobs are likely to go up and up," he says. "If the economy improves, I think we're going to see more companies looking to biocatalysis."

Collaborative successes in the past few years suggest that the field has promise. In one case, Merck teamed up with Codexis, an established biocatalysis company based in Redwood City, California, to develop a new enzymatic synthesis for an existing diabetes drug, sitagliptin. (C. K. Savile *et al. Science* **329**, 305–309; 2010) The traditional production route for the drug involves the use of a rhodium catalyst under high pressure, and the product suffers from rhodium contamination. Investigators streamlined the reaction, reducing by-products and lowering costs by around 10%. The enzymatic route also produced only the desired biologically active form of the molecule, rather than the mixture of forms made by the traditional synthesis. The finding could lay the groundwork for similar syntheses of other drugs.

Codexis is also using its expertise in the fuels sector, and has been working with ►

► Shell since 2006 to develop a liquid transport fuel that is not based on petroleum. The company has grown from a staff of 10 when it was founded in 2002 to around 300 today. Two-thirds of employees work in research and development (R&D), says Dawn Kirkland, director of human resources at Codexis. Some 225 employees are based in the United States, but Codexis also has operations in Singapore and Budapest. The company has grown by 10% every year, and currently has 12 openings for scientists. “We tend to hire both PhD scientists and [master’s or bachelor’s-educated] research associates in the R&D area, and chemical engineers,” says Kirkland. Those with PhDs are most likely to be hired as team or project leaders, she says, but there is no expectation that leadership skills be honed before joining the company — an in-house mentoring system, in which senior scientists counsel new recruits, ensures that rookies get the necessary training.

At Novozymes, a commitment to R&D, including the search for new catalysts, attracts specialists, who tend to stay on. At the moment, about 14% of the company’s revenue is invested in research, says Nickie Spile, vice-president for global R&D. “That’s a lot,” she adds. For comparison, in 2009, the world’s largest chemical company, BASF in Ludwigshafen, Germany, invested 2.7% of its overall revenue in R&D, and Bayer, based in Leverkusen, Germany, invested 8.8%. Opportunities are global: Novozymes has had sites in China and Japan for 20 years, and in 2010 the company hired around 100 people in Europe alone, says Spile.

Other firms working in the field include Johnson Matthey in London, a traditional-catalyst company that in 2010 bought X-Zyme, a biotech based in Düsseldorf, Germany. X-Zyme’s work includes developing biocatalysts that turn ketone and keto-ester molecules into chiral amines, which are used for fine chemicals and drugs. BASF and DSM Pharmaceuticals in Parsippany, New Jersey,

both have a strong interest in biocatalysis; BASF employs about 200 scientists in its R&D unit for “fine chemicals and biocatalysis” (DSM won’t divulge its numbers). DuPont of Wilmington, Delaware, is turning to biocatalysts to make propane-1,3-diol, a building block for polymer and plastics.

A CHEMICAL EDUCATION

A career in biocatalysis should start with a solid academic base, and Europe has some strong groups. The University of Manchester, UK, Delft University of Technology (TU Delft) in the Netherlands and Graz University of Technology in Austria are good places to start, says Nick Turner, director of the Centre of Excellence for Biocatalysis, Biotransformations and Biocatalytic Manufacture (CoEBio3) at the University of Manchester. “You need to get yourself into one of those centres,” he says. Such institutions offer training, have strong links with each other and with industry, and host many of the big academic names in the field.

In 2009, TU Delft started running a two-year master’s degree in biocatalysis. Entrants need a bachelor’s-level education; they learn enzymology, proteomics and biocatalysis, and undertake a design project and a three-month placement in industry. And CoEBio3 received a £3.7-million (US\$6-million) grant from the European Union (EU) in January to coordinate a project that will train PhD students across Europe for industrial work in biocatalysis. The scheme will be run in conjunction with the EU’s Marie Curie training networks, and will initially teach 20 PhD graduates at a time.

Turner says that biocatalysis specialists, even those who concentrate on the biological aspects of the field, need a good grounding in chemistry. There are some biologists at CoEBio3, but it mostly trains chemistry PhD students and postdocs in biotechnical experimental techniques for microbiology, genetic manipulations and protein science. At the moment, around 100 people, including

postdocs and PhDs, are working at the centre, but there is scope to grow, says Turner, who hopes to link a selection of EU-based centres of excellence into a pan-European network.

DIRECTED GROWTH

Rozzell says that much progress in biocatalysis in recent years is thanks to techniques such as directed evolution, in which scientists mutate the genes that code for an enzyme, prompting them to produce thousands of variants. The scientists then pick out the variants that will perform best in the desired industrial process:

those that work well at a specific pH or temperature, or in the presence of a chemical that the enzyme wouldn’t encounter in nature. The technique allows researchers to customize enzymes to fit the ideal conditions for the process.

Directed evolution was pioneered by Frances Arnold, a molecular biologist at the California Institute of Technology in Pasadena, who was Moore’s mentor



Scientists in biocatalysis are helping to create a more sustainable future.

Nickie Spile

during his PhD studies. Moore says that over the past 15 years, his small group at Merck has switched from a laborious microbiological screening technique to exclusive use of directed evolution. “We can do 100 times the work that we used to do,” he says, adding that as the processes become more robust, firms will start to use enzymes to make all sorts of chemicals. Rozzell notes that more and more companies are using directed evolution to produce the enzymes that they then use in large-scale processes.

“It’s a field for the future, quite obviously,” says Ulf Hanefeld, an organic chemist in the biocatalysis group at TU Delft. He cites the manufacturing process for atorvastatin, a cholesterol-lowering drug now made biocatalytically by Pfizer, as a pharmaceutical success for biocatalysis. “There seem to be plenty of jobs,” he says, “but the jobs do require flexibility about where you are based.”

One of the biggest motivations to work in biocatalysis, suggests Spile, is that scientists in the field have a sense of purpose, and that they are involved in an activity that could create a more environmentally-friendly future. For the “sustainably-minded chemist”, biocatalysis can offer a satisfying career track. Moore agrees. “This is one of the places where I really feel I have an opportunity to change the way the world thinks,” he says. ■

Katharine Sanderson is a freelance writer based in Toulouse, France.



The biocatalysis centre at the University of Manchester, UK, is set to lead a Europe-wide training scheme.

UNITED STATES

Mid-career crunch

Some senior scientists feel neglected by the National Institutes of Health's grant formula.

BY KENDALL POWELL

Only a minor spark was needed to set off an online firestorm about the precarious state of US biomedical research funding. In late January on the blog Extramural Nexus, Sally Rockey, deputy director for extramural research at the US National Institutes of Health (NIH), announced the creation of a committee to advise the NIH on the future needs of the biomedical workforce. Daniel Noonan, a molecular biologist at the University of Kentucky in Lexington, wrote in response what he terms a "spontaneous post", outlining what he believes to be problems with current NIH policies that have disproportionately affected funding for mid-career biomedical scientists.

His sentiments struck a chord — resonant to some, and off-key to others. Noonan's post made the e-mail rounds of academic departments and touched off heated online debates about whether the NIH system is 'broken' or headed for disaster, given looming budget concerns. The agency is facing flattened budgets for fiscal years 2011, 2012 and beyond; grant submissions have the lowest success rates in a decade; and policies favour new investigators. Mid-career scientists are under funding pressure, and the situation is creating a growing number of tenured but unfunded professors. Lacking a major NIH grant, these researchers may be forced to shrink their staff numbers, or shift to teaching, administrative or even non-science positions.

In the interest of stretching funds further, Noonan advised the NIH to delay large initiatives such as building the National Center for Advancing Translational Science; limit individual investigators to three grants or US\$1 million a year; implement a formula that gives investigators with multiple grants less per grant for indirect costs from university overheads; factor in an investigator's non-NIH funding when deciding whether to provide a grant; and limit or eliminate funds for construction projects.

Applications for NIH grants are scored for scientific merit; for investigators beyond the early stages of their careers, only those scoring in roughly the top 10% get funded. In 2010, success rates for R01 grants — the NIH's primary grant for individual investigators — were

around 15–25%. "If you lose that one grant-renewal opportunity, it's hard to recover in this day and age," says Noonan. The pool of money dedicated to investigator-initiated grants has shrunk, he says, and with limited state and federal budgets, the NIH needs to find ways to generate money from within.

Rockey has pledged to forward the online discussions to the workforce advisory committee, which will be chaired by Shirley Tilghman, a molecular biologist and president of Princeton University in New Jersey. But Rockey says that no decisions have been made to cap applications from individuals. "There are a lot of

individuals has received some attention. Several investigators receive multiple NIH grants: a 2008 analysis by *Nature* found 200 scientists who held 6 or more (see *Nature* 452, 258–259; 2008). And the ASBMB calculates that in fiscal year 2009, 1,600 scientists each received \$1 million or more from the NIH. Rockey points out that the National Institute of General Medical Sciences already has a capping policy — an investigator already receiving \$750,000 or more from any source must justify new grant submissions.

But she emphasizes that such multiple-grant holders are rare. "Contrary to popular belief, the average NIH-funded scientist holds 1.4 grants at any one time," she says. "So there is not a huge cadre of people who have eight, nine or ten grants." But she concedes that compared with a decade ago, more researchers are fighting over the same sized slice of pie, as a result of budget doubling that has now levelled off.

"It's a difficulty for just about everybody — early-, mid- or late-career — in sustaining NIH funding," says Rockey, noting that scientists' frustrations stem from "a lot of pent-up good science going on that we are unable to fund." Tilghman's workforce advisory committee will try to determine the size and composition of the biomedical workforce that the NIH can support.

Not everyone thinks that the NIH needs to reconsider how grant funding is apportioned. Several scientist bloggers believe that Noonan's comments imply that scientists should have access to NIH resources regardless of ability or outcomes; they counter that meritocracy should rule. One contributor wrote that with budgets shrinking, researchers really should be concentrating on communicating the value of research to the public.

Rockey advises mid-career scientists facing an R01 renewal to consider a no-cost extension (stretching out existing grant funds) for another year to gather data or publish results; or to try making contingency plans, such as seeking bridge funding from their institutions. "Your reviewer is taking into account what you have already accomplished, so be sure to highlight how well your research is going and the strengths of your research team," she says. Rockey's top recommendation, she says, is to seek advice from the relevant grant-review programme officer. ■



different thoughts about how one might go about this," she says. "We want to have data and facts and information before we resort to any social engineering of the workforce."

Ideas for reform extend beyond unhappy individual investigators. The 12,000-member American Society for Biochemistry and Molecular Biology (ASBMB) in Bethesda, Maryland, has proposed some recommendations, says Benjamin Corb, the society's director of public affairs. The ASBMB seeks a cap on the funding going to any one person, and suggests that money be redirected to the R01 pool from large initiatives that have not made medical breakthroughs, such as the Genome-Wide Association Studies programme and the Protein Structure Initiative. The ASBMB also proposes that the NIH adopt a sliding scale, to partially fund lower-scoring but meritorious grants. The society presented its recommendations to six institute directors and Lawrence Tabak, the NIH's deputy director, on 14 March.

The idea of capping the number of awards to

ENTANGLEMENT

A world of possibilities.

BY MARISSA LINGEN

Ever since the quantum universes reunited, I haven't been able to find my favourite pen.

It seems like such a small thing, but there it is, or rather, there it isn't: a black fountain pen with gold in the nib. I had black ink in it at the time, the only black ink I've ever found that was really dark enough. The grip was smoothed to my hand. I'm old-fashioned, I guess, but little things like that matter to me.

I remember losing it at the Midwest Medieval History Conference, just as I remember putting it neatly away in the drawer when I got home. This quantum re-entanglement: it's harder than it sounds.

I firmly remember my wife always being alive, except that I also firmly remember that she had a rare kidney infection that killed her two years ago. You would think this would be a bigger deal than the pen, and it is, but there's nothing I can do about it. I keep thinking the pen must be somewhere, and I know where Sara is; she's right here. Right where she's always been. Except for the two years I spent grieving for her every day and night.

Sara has memories of dying, too, which is as hard as you might think on her. Maybe. She says it could be worse. She says she is less afraid.

I am more afraid.

I am afraid of what other worlds are going to come in on us in a tangle.

I am afraid of my job going wrong, although it hasn't; they say that these worlds were able to re-entangle because the large-scale events stayed so similar. But I'm afraid next time we won't be so lucky.

I am afraid of losing my wife again.

As a historian, I study the ins and outs of what makes each path, what contributes to each decision, and I want them to make sense. I want them to be inevitable. I want them to follow patterns that I can find, so that I can say: "If the French had done this, then the Burgundians would have done that — *but they didn't*. If the Spanish had responded thusly — *but no*."

But this, this re-entanglement, it says yes.

It says they did. It says that the patterns are all followed. Not here. But somewhere. And we could get back there.

Perhaps if we got back there, I could find my pen.

Perhaps if we got back there, I would lose my wife.

The rest of the world — the rest of these united worlds — I don't know how to say it, even. The people around me seem to be navigating by touch, by not thinking of what they have gained, or what they have lost, in this reunification.

Last month I flew to Germany. I asked people on the street whether it was like their other reunification. Most of the people I asked were my own age, too young to really remember what it was like. I identified myself as a historian, and they sent me trundling off to an old folks' care home, where people would care about history. I should be used to this, in any world, in any reality, but the truth is that my field is obscure enough that even the elderly are presumed not to mind about it.

The elderly were more like me — more disturbed, more upset. I felt fussy among them, rebellious against my own interest, but I had to ask again, about reunification.

"The Communists did not take my favourite slippers," said one woman in impeccable

English, although I had asked her in German, "and they did not undo the dishes I had already done. For my part, I prefer the Communists."

"Former Communists," muttered her roommate.

The first woman ignored her. "We are Germans. We understand quantum mechanics. We understand that we do *not* understand quantum mechanics, and that is understanding quantum mechanics. Now they tell us that there is re-entanglement and that is why the trains are here and gone and cats we remember burying are still waking us to ask for food. Very well."

"Very well?" I demanded. "Very well?"

"It is not a cat for you."

"My wife," I wanted to say, but instead, I said: "You haven't seen a black fountain pen? I think I've lost it."

"Young man, they never should have told us. They should have let us think we were confused. We would have each made up stories and smoothed it over for ourselves. Now we can't go back."

"I have to go back," I said, but I didn't see how, except literally, and the literal had never satisfied me.

I took a taxi back from the airport.

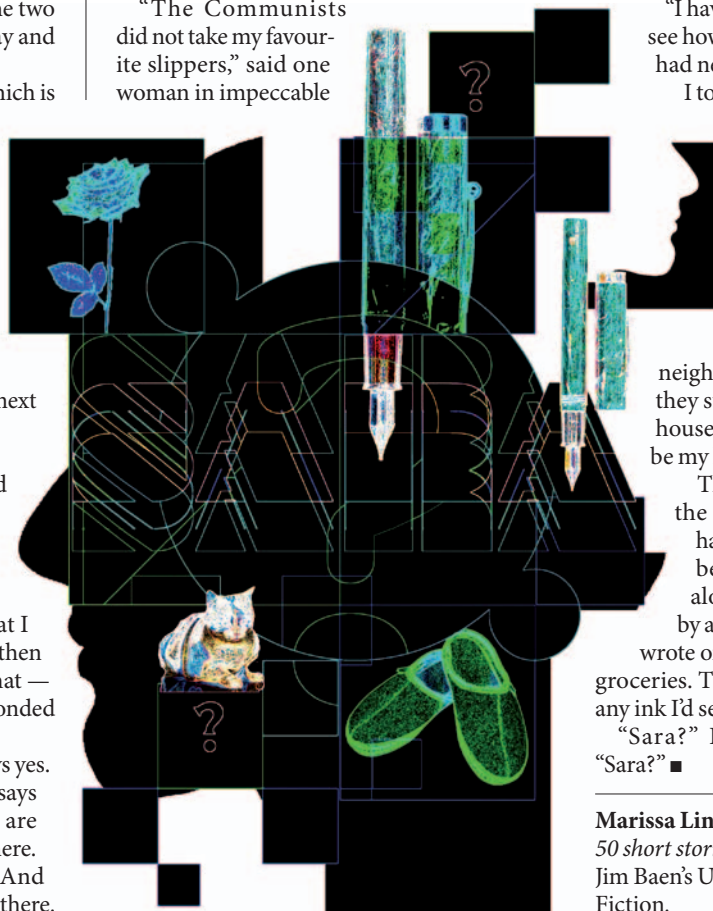
I couldn't tell whether I was jet-lagged or something worse. Quantum mechanics would be the death of me, I felt sure. I was second-guessing every billboard: had they changed it while I was gone, or had it always been like that? Had the

neighbour pruned her roses, or were they supposed to be hydrangeas? My house was my house. Surely it would be my house.

There was a red fountain pen in the middle of the kitchen desk. I had never seen it before. I remembered buying it on a trip I took alone, after my wife Sara was hit by a car and died. I uncapped it and wrote on the notepad we kept there for groceries. The ink was black, blacker than any ink I'd seen.

"Sara?" I called into the bedroom. "Sara?" ■

Marissa Lingen has published more than 50 short stories in venues such as *Analog*, *Jim Baen's Universe* and *Aeon Speculative Fiction*.



TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA

Rachel Litman Flynn¹, Richard C. Centore^{1*}, Roderick J. O'Sullivan^{3*}, Rekha Rai^{4*}, Alice Tse¹, Zhou Songyang⁵, Sandy Chang⁴, Jan Karlseder³ & Lee Zou^{1,2}

Maintenance of telomeres requires both DNA replication and telomere 'capping' by shelterin. These two processes use two single-stranded DNA (ssDNA)-binding proteins, replication protein A (RPA) and protection of telomeres 1 (POT1). Although RPA and POT1 each have a critical role at telomeres, how they function in concert is not clear. POT1 ablation leads to activation of the ataxia telangiectasia and Rad3-related (ATR) checkpoint kinase at telomeres^{1,2}, suggesting that POT1 antagonizes RPA binding to telomeric ssDNA. Unexpectedly, we found that purified POT1 and its functional partner TPP1 are unable to prevent RPA binding to telomeric ssDNA efficiently. In cell extracts, we identified a novel activity that specifically displaces RPA, but not POT1, from telomeric ssDNA. Using purified protein, here we show that the heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) recapitulates the RPA displacing activity. The RPA displacing activity is inhibited by the telomeric repeat-containing RNA (TERRA) in early S phase, but is then unleashed in late S phase when TERRA levels decline at telomeres³. Interestingly, TERRA also promotes POT1 binding to telomeric ssDNA by removing hnRNPA1, suggesting that the reaccumulation of TERRA after S phase helps to complete the RPA-to-POT1 switch on telomeric ssDNA. Together, our data suggest that hnRNPA1, TERRA and POT1 act in concert to displace RPA from telomeric ssDNA after DNA replication, and promote telomere capping to preserve genomic integrity.

RPA binds ssDNA in a non-sequence specific manner⁴, whereas POT1 specifically recognizes ssDNA consisting of the telomeric repeats⁵. RPA plays a key role in DNA replication and activation of the ATR checkpoint⁶, and POT1 suppresses ATR activation at telomeres^{1,2} (Supplementary Fig. 1). In both yeast and humans, RPA associates with telomeres during S phase of the cell cycle^{7–9}, and is implicated in telomere maintenance^{10–12}. Furthermore, ATR transiently associates with telomeres and suppresses telomere instability^{7,10,13}. These findings raise the question of how the bindings of POT1 and RPA to telomeric ssDNA are orchestrated and, furthermore, how the interplay between POT1 and RPA affects DNA replication and ATR activation at telomeres.

Double-stranded DNA (dsDNA) with ssDNA overhangs activates ATR in human cell extracts¹⁴. To investigate how ATR activation is suppressed at telomeres, we tested whether telomeric ssDNA overhangs affect ATR activation in this assay. Resected dsDNA of random sequences, but not resected telomeric dsDNA, efficiently induced the phosphorylation of RPA2 by ATR (Supplementary Fig. 2)¹⁴, suggesting that telomeric ssDNA overhangs do not support efficient ATR activation in cell extracts.

The absence of ATR activation by telomeric ssDNA suggests that POT1 may prevent RPA binding to telomeric ssDNA². POT1 and TPP1 function as heterodimers in cells, and the complex binds to telomeric ssDNA more efficiently than POT1 alone^{15,16}. In gel-shift assays, the POT1–TPP1 complexes purified from insect or human cells

and the RPA purified from *Escherichia coli* efficiently bound to a telomeric ssDNA probe (Fig. 1a and Supplementary Fig. 3a, b). POT1–TPP1 exhibited lower affinity for telomeric ssDNA than RPA (Supplementary Fig. 3a). When POT1–TPP1 and RPA were co-incubated with the probe, the RPA–ssDNA complex was readily detected, whereas no POT1-containing complexes were observed (Fig. 1a and Supplementary Fig. 3b). In pull-down assays using biotinylated telomeric ssDNA (ssTEL), RPA also outcompeted POT1–TPP1 for binding to ssTEL (Fig. 1b and Supplementary Fig. 3c). Thus, RPA, which is more abundant than POT1–TPP1 in cells^{4,17}, outcompetes POT1–TPP1 for binding to telomeric ssDNA when present at similar concentrations as POT1–TPP1. The *E. coli* ssDNA-binding protein only modestly reduced POT1 binding to ssTEL (Supplementary Fig. 3c), suggesting that the ability to outcompete POT1–TPP1 is unique to RPA.

The ability of RPA to outcompete POT1–TPP1 raises the question of how ATR activation is suppressed in cell extracts. Purified RPA bound to ssTEL and mutated telomeric repeats (ssMUT) efficiently (Fig. 1c). In stark contrast to purified RPA, the endogenous RPA in HeLa whole-cell extracts (WCEs) was largely excluded from ssTEL; however, it still associated with ssMUT (Fig. 1c). The sequence-specific exclusion of RPA from ssTEL in WCEs suggests that RPA may be outcompeted by other proteins or actively displaced from telomeric ssDNA.

To assess if RPA is actively displaced from ssTEL, we pre-coated ssTEL and ssMUT with RPA then incubated them in extracts. The levels of RPA on ssTEL gradually declined with increasing concentrations of WCEs from HeLa, HEK293E, U2OS and MEF cells (Fig. 1d and Supplementary Fig. 4a). In addition, HeLa nuclear extracts, but not the cytoplasmic extracts, efficiently displaced RPA from ssTEL (Supplementary Fig. 4b). In marked contrast to the RPA on ssTEL, the RPA bound to ssMUT remained constant regardless of WCE concentrations (Fig. 1d). When POT1-coated ssTEL was incubated in extracts, POT1 remained stably bound to ssTEL even in high concentrations of WCEs (Fig. 1e). Furthermore, RPA was rapidly displaced from ssTEL within 5 min, whereas no POT1 was displaced after 60 min (Supplementary Fig. 4c). Thus, the activity that displaces RPA from telomeric ssDNA is sequence-specific, protein-specific and localized within the nucleus.

The specific displacement of RPA, but not POT1, from telomeric ssDNA prompted us to test if POT1 is the RPA displacing factor. When incubated with RPA–ssTEL, POT1–TPP1 did not significantly reduce the levels of ssTEL-bound RPA (Supplementary Fig. 5a). To identify the RPA displacing factors, we sought to capture the RPA displacing activity from extracts using RPA–ssTEL as bait. The proteins captured and eluted from RPA–ssTEL, but not RPA–ssMUT, recapitulated the RPA displacing activity (Fig. 2a). Mass spectrometry analysis of the proteins specifically captured by RPA–ssTEL identified hnRNPA1

¹Massachusetts General Hospital Cancer Center, Harvard Medical School, Charlestown, Massachusetts 02129, USA. ²Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA. ³Molecular and Cellular Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA. ⁴Department of Laboratory Medicine, Yale University School of Medicine, New Haven, Connecticut 06510, USA. ⁵Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas 77030, USA.

*These authors contribute equally to this work.

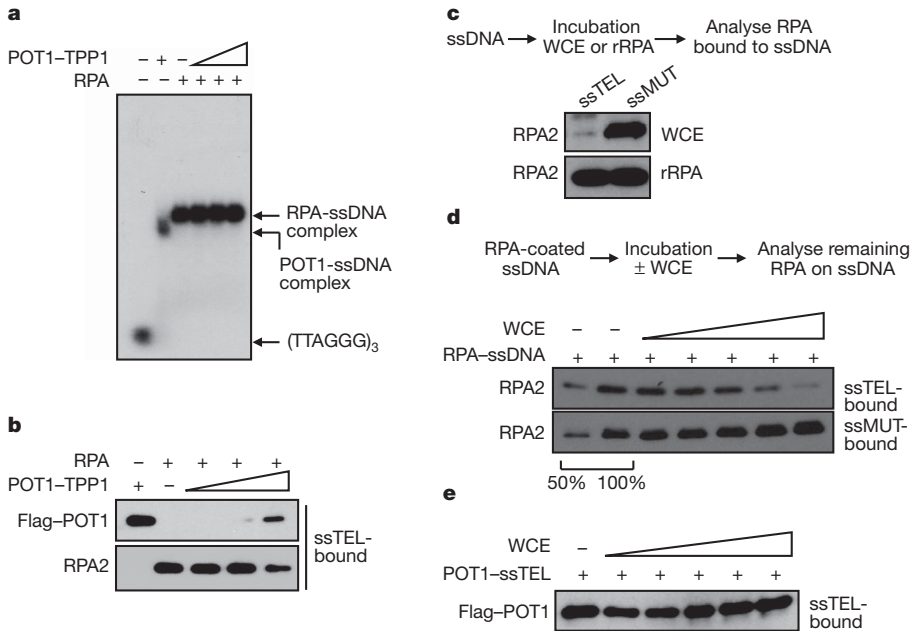


Figure 1 | A novel telomere-specific RPA displacing activity in human cell extracts. **a**, POT1–TPP1 (60 nM; purified from insect cells), RPA (60 nM) and mixtures of POT1–TPP1 and RPA (60, 120, 180 nM of POT1–TPP1 mixed with 60 nM of RPA) were incubated with 20 nM of the ssDNA probe and analysed by gel-shift. **b**, POT1–TPP1 (2.4 nM), RPA (2.4 nM) and mixtures of POT1–TPP1 and RPA (2.4, 4.8, 7.2 nM of POT1–TPP1 mixed with 2.4 nM of RPA) were incubated with 0.8 nM of biotinylated ssTEL (TTTAGG₃). The

and hnRNP A2/B1 (Supplementary Fig. 5b, c), both of which are known to bind telomeric ssDNA^{18,19}. The presence of hnRNP A1 and A2/B1 in the eluted fraction with RPA displacing activity was confirmed by western blot (Fig. 2b). Moreover, hnRNP A1 and A2/B1 gradually bound to ssTEL as RPA was displaced in WCEs (Fig. 2c). These results suggest that hnRNP A1 and A2/B1 may play a role in RPA displacement.

hnRNPA1 has been implicated in telomere maintenance^{20,21}. Extracts from hnRNPA1 knockdown cells exhibited reduced activity

proteins bound to ssTEL were retrieved by streptavidin beads and analysed by western blot. **c**, Biotinylated ssTEL or ssMUT ((TTTGGC)₈) was incubated with WCEs or recombinant RPA (rRPA). **d**, ssTEL or ssMUT pre-coated with RPA was incubated with increasing concentrations of HeLa WCEs (0.08, 0.19, 0.36, 0.8, 1.3 $\mu\text{g } \mu\text{l}^{-1}$). The RPA2 remaining on ssTEL was analysed as in **b**. **e**, ssTEL pre-coated with POT1 was incubated with increasing concentrations of HeLa WCEs (0.07, 0.18, 0.33, 0.66, 1.3 $\mu\text{g } \mu\text{l}^{-1}$).

in RPA displacement (Supplementary Fig. 6a). Purified hnRNPA1 efficiently displaces RPA from ssTEL, but not ssMUT (Fig. 2d). Furthermore, hnRNPA1 did not displace POT1 from ssTEL (Fig. 2e). hnRNPA1 only displaces RPA from ssTEL containing four or more telomeric repeats (Supplementary Fig. 6b), indicating that a DNA length-dependent binding mode of hnRNPA1 may be needed to displace RPA²². Given that hnRNPA1 and A2/B1 are highly homologous in the RRM domains that bind telomeric ssDNA, both of these hnRNPs may contribute to RPA displacement.

hnRNPA1 not only binds telomeric ssDNA but also TERRA^{23–26}. To test if TERRA affects the ability of hnRNPA1 to displace RPA from ssTEL, we added increasing concentrations of TERRA or control RNA to nuclear extracts, then incubated the extracts with RPA–ssTEL. RPA displacement was virtually abolished by TERRA, but not control RNA (Fig. 3a). The RPA displacing activity captured by RPA–ssTEL was also specifically inhibited by TERRA (Fig. 3b). Furthermore, the ability of purified hnRNPA1 to bind ssTEL and to displace RPA from ssTEL was specifically inhibited by TERRA (Fig. 3c and Supplementary Fig. 6c). Thus, TERRA is a potent inhibitor of the RPA displacing activity of hnRNPA1.

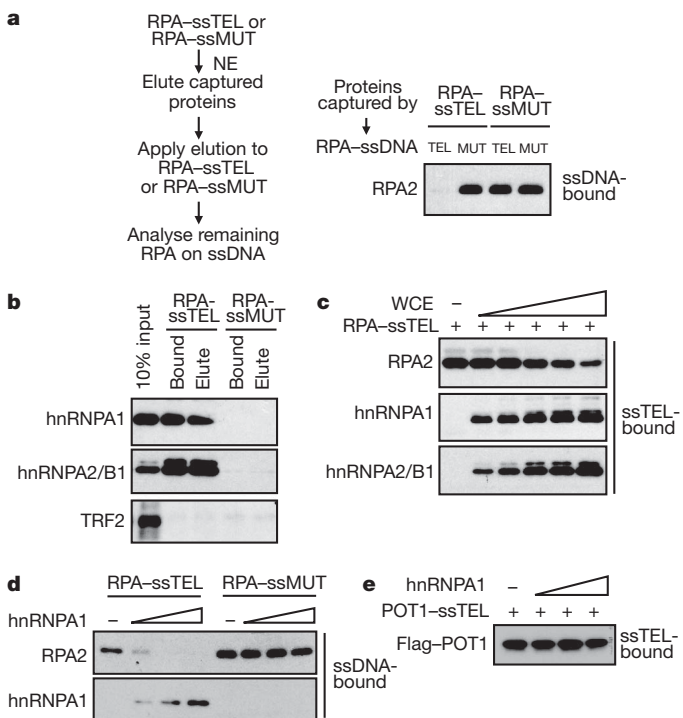


Figure 2 | RPA displacement by hnRNP1. **a**, ssTEL and ssMUT pre-coated with RPA were incubated with nuclear extracts (NE). After the incubation, the proteins bound to DNA were retrieved, eluted and applied to RPA-coated ssTEL or ssMUT (see Supplementary Methods). After the second incubation, the remaining RPA2 on DNA was analysed by western blot. **b**, Proteins captured by RPA–ssTEL or RPA–ssMUT and eluted by salt were analysed by western blot using antibodies to hnRNP1, hnRNP2/B1 and TRF2. **c**, RPA-coated ssTEL (0.8 nM) was incubated with increasing concentrations of WCEs (0.06, 0.24, 0.96 $\mu\text{g } \mu\text{l}^{-1}$). The hnRNP1 and hnRNP2/B1 bound to DNA and the remaining RPA2 on DNA were analysed by western blot. **d**, RPA-coated ssTEL or ssMUT (0.8 nM) was incubated with increasing concentrations of purified hnRNP1 (2.4, 4.8, 7.2 nM). The remaining RPA2 on DNA was analysed as in **a**. **e**, POT1-coated ssTEL (0.8 nM) was incubated with increasing concentrations of purified hnRNP1 (2.4, 4.8, 7.2 nM).

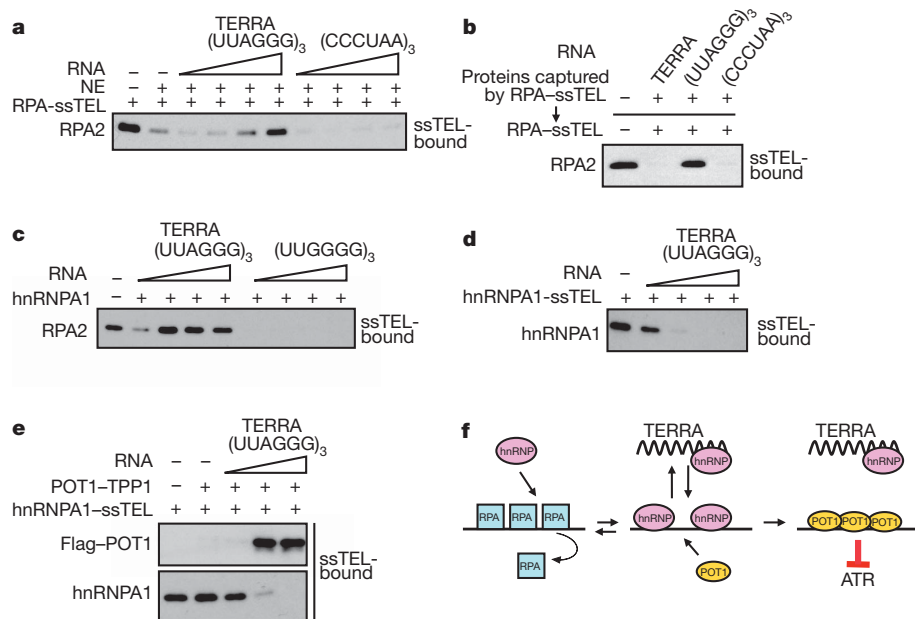


Figure 3 | Regulation of RPA displacement by TERRA. **a**, Nuclear extracts ($34 \text{ ng } \mu\text{l}^{-1}$) were treated with increasing concentrations (2, 4, 10, 20 nM) of TERRA (UUAGGG)₃, control RNA (CCCUGA)₃ or mock treated. The treated nuclear extracts were then incubated with RPA-coated ssTEL (2 nM), and the remaining RPA2 on ssTEL was analysed after the incubation. **b**, The RPA displacing factors were captured with RPA-ssTEL as in Fig. 2a. The elution was incubated with TERRA or control RNA, then applied to RPA-ssTEL. **c**, Purified hnRNPA1 (4.8 nM) was incubated with increasing concentrations of

TERRA or control RNA (2, 4, 10, 20 nM), then incubated with RPA-ssTEL (0.8 nM). **d**, hnRNPA1-coated ssTEL (0.8 nM) was incubated with increasing concentrations of TERRA (2, 20, 200, 2,000 nM). The remaining hnRNPA1 on ssTEL was analysed by western blot. **e**, hnRNPA1-coated ssTEL (0.8 nM) was incubated with increasing concentrations of TERRA (2, 20, 200 nM) in the presence of POT1-TPP1 (2.4 nM). The hnRNPA1 and POT1 on ssTEL were analysed by western blot. **f**, A model for RPA displacement.

If hnRNPA1 displaces RPA from telomeric ssDNA, how can POT1 bind to telomeric ssDNA? Given that hnRNPA1 has affinity for both telomeric ssDNA and TERRA, the presence of TERRA at telomeres may promote the dissociation of hnRNPA1 from telomeric ssDNA. Indeed, when hnRNPA1-coated ssTEL was incubated with TERRA, hnRNPA1 was stripped from ssTEL (Fig. 3d), showing that hnRNPA1 binds telomeric ssDNA dynamically. Furthermore, when hnRNPA1-ssTEL was incubated with TERRA and POT1-TPP1, POT1 efficiently bound to ssTEL as hnRNPA1 was removed by TERRA (Fig. 3e).

The *in vitro* results above suggest that the initial displacement of RPA from telomeric ssDNA may be performed by hnRNPs when TERRA levels are low at telomeres (Fig. 3f). However, if TERRA levels rise at telomeres, hnRNPA1 may shuttle between telomeric ssDNA and TERRA dynamically. In this situation, both RPA and POT1 may have the chance to bind telomeric ssDNA. Because hnRNPA1 only displaces RPA, but not POT1, this dynamic process will eventually promote POT1 occupancy at telomeric ssDNA.

This model raises the possibility that the RPA displacing activity may be regulated by TERRA during the cell cycle. To test this, we generated WCEs from cells in G1, early S, late S and M phases of the cell cycle. RPA was more efficiently displaced in the late S- and M-phase extracts than in the G1- and early S-phase extracts (Fig. 4a and Supplementary Fig. 7a, b). Thus, the RPA displacing activity is low in G1 and early S phase, but upregulated in late S phase.

If TERRA inhibits the RPA displacing activity, its levels should inversely correlate with the activity. Furthermore, removal of TERRA in early S phase should alleviate the inhibition. Indeed, a recent study showed that TERRA levels significantly decrease in late S phase and increase again after S phase³. Consistently, telomeric TERRA foci declined as cells progressed from early to late S phase (Fig. 4b and Supplementary Fig. 7c, d). In addition, RNase A treatment of early S-phase extracts significantly enhanced the RPA displacing activity (Supplementary Fig. 7e). Together, these results suggest that TERRA inhibits RPA displacement in early S phase, and its decline in late S phase may provide a window for RPA displacement.

The model above also predicts that hnRNPs are necessary for RPA displacement from telomeres. Depletion of hnRNPA1 using two independent short interfering RNA (siRNAs) significantly increased the fraction of cells displaying RPA foci (Fig. 4c and Supplementary Fig. 8a–d). Notably, a fraction of the RPA foci in hnRNPA1 knockdown cells closely associated with TRF2 foci. Furthermore, increased RPA binding at telomeres was detected in hnRNPA1 knockdown cells by chromatin immunoprecipitation (Fig. 4d). In synchronized hnRNPA1 knockdown cells, RPA binding to telomeres was enhanced in early S phase (Supplementary Fig. 9a, b), indicating that even during this period some hnRNPA1 remains free from TERRA and limits RPA binding to telomeres⁹. In late S/G2, RPA still declined at telomeres in hnRNPA1 knockdown cells, possibly owing to the redundancy among hnRNPs.

If the displacement of RPA by hnRNPA1 is a prerequisite for POT1 binding, POT1 should be needed for RPA exclusion after late S phase. To assess this possibility, we treated cells with POT1 siRNA and synchronized the cells with thymidine as POT1 levels declined (Supplementary Fig. 10a, b). After POT1 knockdown cells and control cells were synchronously released, RPA foci appeared in both cell populations (Fig. 4e). As control cells entered G2, RPA foci rapidly declined. In contrast, the POT1 knockdown cells containing RPA foci that colocalized with TRF2 continued to increase. Concomitantly, modest Chk1 phosphorylation was detected in POT1 knockdown cells (Supplementary Fig. 10c). Thus, reduction of POT1 compromises the exclusion of RPA from telomeres after replication²⁷.

During early to middle S phase, TERRA sequesters hnRNPs and allows RPA to bind telomeric ssDNA at replication forks or telomere ends (Supplementary Fig. 1). When TERRA levels decline in late S phase, hnRNPs are unleashed to displace RPA from telomeric ssDNA. The dynamic binding of hnRNPs to telomeric ssDNA is gradually antagonized by TERRA when TERRA reaccumulates at telomeres, providing a window for both RPA and POT1 to bind. Because only POT1, but not RPA, binds to telomeric ssDNA irreversibly in the presence of hnRNPs, this dynamic process favours the formation of

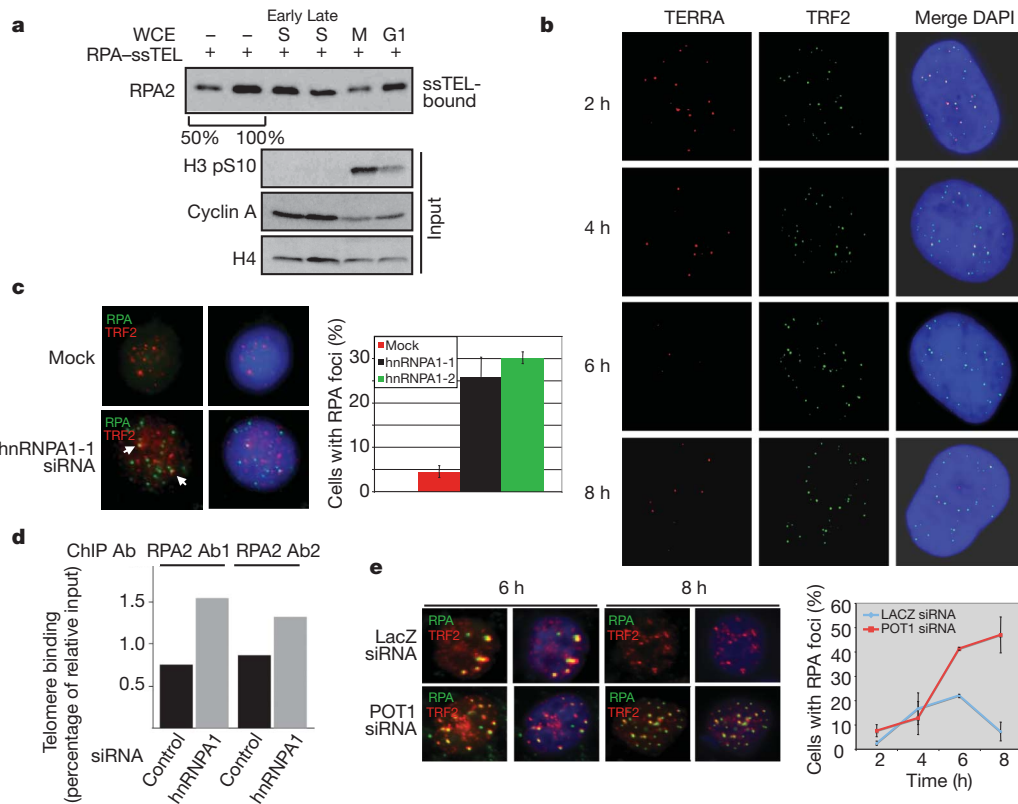


Figure 4 | hnRNPA1 and POT1 suppress the accumulation of RPA at telomeres. **a**, RPA-coated ssTEL was incubated with WCEs from cells in G1, early S, late S and M phases of the cell cycle (see Supplementary Methods). The remaining RPA2 on ssTEL was analysed after incubation. Cyclin A and phospho-histone H3 serve as cell-cycle markers, and histone H4 as a loading control. **b**, TERRA was analysed by RNA fluorescence *in situ* hybridization in HeLa cells after thymidine release. TERRA serves as a marker of telomeres. **c**, HeLa cells were treated with hnRNPA1 siRNA or mock treated, then immunostained with antibodies to RPA2 and TRF2 (left panel). The cells with

RPA foci (>5) were quantified (right panel). Mean \pm s.d., $n = 3$ for mock and sihnRNPA1-1, $n = 2$ for sihnRNPA1-2. **d**, Chromatin immunoprecipitation (ChIP) of RPA was performed with two different RPA2 antibodies (Ab). The association of RPA with telomeres was analysed by dot blot using a telomere probe and quantified. **e**, HeLa cells transfected with POT1 or LacZ siRNA were released from a thymidine block. At the indicated times, the G2/M population was determined by FACS (Supplementary Fig. 10c). Cells were immunostained for RPA2 and TRF2 (left panel). The cells with RPA foci (>5) were quantified (right panel). Mean \pm s.d., $n = 3$.

POT1-coated telomeric ssDNA. Unlike RPA, POT1 kinks telomeric ssDNA and induces its self-recognition^{28,29}. These unique properties of POT1 may confer resistance to hnRNP-mediated displacement. The cell-cycle-regulated RPA displacement may allow RPA to transiently associate with telomeric ssDNA during replication, and prevent persistent ATR activation at telomeres after S phase. Once coated by POT1, telomeric ssDNA may remain capped until the arrival of replication forks in the next S phase. Together, TERRA and hnRNPs orchestrate a cell-cycle-regulated RPA-to-POT1 switch on telomeric ssDNA, ensuring orderly telomere replication and capping.

METHODS SUMMARY

To analyse the bindings of RPA and POT1–TPP1 to ssDNA, biotinylated ssDNA was attached to streptavidin-coated magnetic beads. Biotinylated ssDNA (1 pmol) was incubated with purified protein in 500 μ l of binding buffer (10 mM Tris–HCl (pH 7.5), 100 mM NaCl, 10 μ g ml^{−1} BSA, 10% glycerol, 0.05% NP-40).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 August; accepted 22 December 2010.

Published online 13 March 2011.

- Guo, X. *et al.* Dysfunctional telomeres activate an ATM–ATR-dependent DNA damage response to suppress tumorigenesis. *EMBO J.* **26**, 4709–4719 (2007).
- Denchi, E. L. & de Lange, T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* **448**, 1068–1071 (2007).
- Porro, A., Feuerhahn, S., Reichenbach, P. & Lingner, J. Molecular dissection of TERRA biogenesis unveils the presence of distinct and multiple regulatory pathways. *Mol. Cell Biol.* **30**, 4808–4817 (2010).

- Wold, M. S. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* **66**, 61–92 (1997).
- Baumann, P. & Cech, T. R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292**, 1171–1175 (2001).
- Zou, L. & Elledge, S. J. Sensing DNA damage through ATRIP recognition of RPA–ssDNA complexes. *Science* **300**, 1542–1548 (2003).
- Verdun, R. E. & Karlseder, J. The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres. *Cell* **127**, 709–720 (2006).
- Moser, B. A. *et al.* Differential arrival of leading and lagging strand DNA polymerases at fission yeast telomeres. *EMBO J.* **28**, 810–820 (2009).
- McGee, J. S. *et al.* Reduced Rif2 and lack of Mec1 target short telomeres for elongation rather than double-strand break repair. *Nature Struct. Mol. Biol.* **17**, 1438–1445 (2010).
- Verdun, R. E., Crabbe, L., Haggblom, C. & Karlseder, J. Functional human telomeres are recognized as DNA damage in G2 of the cell cycle. *Mol. Cell* **20**, 551–561 (2005).
- Kibe, T., Ono, Y., Sato, K. & Ueno, M. Fission yeast Taz1 and RPA are synergistically required to prevent rapid telomere loss. *Mol. Biol. Cell* **18**, 2378–2387 (2007).
- Schramke, V. *et al.* RPA regulates telomerase action by providing Est1p access to chromosome ends. *Nature Genet.* **36**, 46–54 (2004).
- McNees, C. J. *et al.* ATR suppresses telomere fragility and recombination but is dispensable for elongation of short telomeres by telomerase. *J. Cell Biol.* **188**, 639–652 (2010).
- Shiotani, B. & Zou, L. Single-stranded DNA orchestrates an ATM-to-ATR switch at DNA breaks. *Mol. Cell* **33**, 547–558 (2009).
- Wang, F. *et al.* The POT1–TPP1 telomere complex is a telomerase processivity factor. *Nature* **445**, 506–510 (2007).
- Xin, H. *et al.* TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. *Nature* **445**, 559–562 (2007).
- Takai, K. K., Hooper, S., Blackwood, S., Gandhi, R. & de Lange, T. *In vivo* stoichiometry of shelterin components. *J. Biol. Chem.* **285**, 1457–1467 (2010).
- Ishikawa, F., Matunis, M. J., Dreyfuss, G. & Cech, T. R. Nuclear proteins that bind the pre-mRNA 3' splice site sequence (UUAG/G) and the human telomeric DNA sequence d(TTAGGG)_n. *Mol. Cell Biol.* **13**, 4301–4310 (1993).

19. McKay, S. J. & Cooke, H. hnRNP A2/B1 binds specifically to single stranded vertebrate telomeric repeat TTAGGGn. *Nucleic Acids Res.* **20**, 6461–6464 (1992).
20. LaBranche, H. *et al.* Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. *Nature Genet.* **19**, 199–202 (1998).
21. Zhang, Q. S., Manche, L., Xu, R. M. & Krainer, A. R. hnRNP A1 associates with telomere ends and stimulates telomerase activity. *RNA* **12**, 1116–1128 (2006).
22. Ding, J. *et al.* Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev.* **13**, 1102–1115 (1999).
23. Deng, Z., Norseen, J., Wiedmer, A., Riethman, H. & Lieberman, P. M. TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol. Cell* **35**, 403–413 (2009).
24. de Silanes, I. L., d'Alcontres, M. S. & Blasco, M. A. TERRA transcripts are bound by a complex array of RNA-binding proteins. *Nature Commun.* **1**, 1–9 (2010).
25. Azzalin, C. M., Reichenbach, P., Khoraiuli, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798–801 (2007).
26. Redon, S., Reichenbach, P. & Lingner, J. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res.* **38**, 5797–5806 (2010).
27. Gong, Y. & de Lange, T. A. Shld1-controlled POT1a provides support for repression of ATR signaling at telomeres through RPA exclusion. *Mol. Cell* **40**, 377–387 (2010).
28. Lei, M., Podell, E. R. & Cech, T. R. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nature Struct. Mol. Biol.* **11**, 1223–1229 (2004).
29. Bochkarev, A., Pfuetzner, R. A., Edwards, A. M. & Frappier, L. Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**, 176–181 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. de Lange, A. Krainer, B. Chabot and M. Wold for reagents, and members of the Zou laboratory for discussion. L.Z. is an Ellison New Scholar on Aging. R.L.F. is supported by National Institutes of Health (NIH) fellowship 5T32CA009216-28 and American Cancer Society fellowship 0902501. R.C.C. is supported by NIH fellowship F32-GM089150. R.J.O. is supported by the George E. Hewitt Foundation for Medical Research. This work is supported by a Welch Foundation grant (Q-1673) to Z.S., an ACS grant (RSG-08-297) to L.Z., and NIH grants CA133249 to Z.S., CA129037 to S.C., GM06525 and AG025837 to J.K. and GM076388 to L.Z.

Author Contributions R.L.F. and L.Z. conceived the project. R.L.F., R.C.C., R.J.O., R.R., A.T. and L.Z. performed the experiments. Z.S. contributed the POT1–TPP1 complex purified from insect cells. R.R. and S.C. performed the combined RNA-fluorescence *in situ* hybridization and immunostaining analysis. R.J.O. and J.K. performed the RPA chromatin immunoprecipitation and northern blot. R.L.F. and L.Z. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.Z. (zou.lee@mgh.harvard.edu).

METHODS

ATR activation. To generate the 800-base-pair (bp) telomeric dsDNA fragment, the pSTY11 plasmid (a gift from T. de Lange) was digested with EcoRI and the excised fragment was gel purified. The 800-bp random sequence dsDNA was generated by PCR and column purified. These dsDNA fragments were incubated with T7 exonuclease for 15 s at room temperature (approximately 23 °C) and flash frozen in an ethanol-dry-ice bath. T7 was inactivated by subsequent incubation at 70 °C for 20 min and DNA fragments were separated on 2% agarose gel to confirm equal resection. The resected DNA fragments were incubated with nuclear extract as previously described¹⁴. To monitor specifically the phosphorylation of RPA2 by ATR and eliminate the contributions of ataxia telangiectasia mutated (ATM) and DNA-PK to RPA2 phosphorylation, nuclear extracts were pre-treated with 20 μ M KU55933 and NU7026 inhibitors for 15 min at 4 °C. The extracts were mixed with the DNA fragments, incubated for 15 min at 37 °C, and RPA phosphorylation was analysed by western blot.

Protein purification. The POT1–TPP1 complex was either purified from baculovirus-infected Sf9 cells as previously described¹⁶, or purified from HEK293E cells as follows. The pCL-Flag–POT1 and pCL-Flag–TPP1 vectors¹⁶ were individually transfected or co-transfected into HEK293E cells. The cells were collected after 72 h and lysed in the NETN buffer (100 mM NaCl, 1 mM EDTA, 20 mM Tris-HCl (pH 8.0), 0.5% NP-40 and protease inhibitors), sonicated and cleared by centrifugation (10,000g for 10 min). The cleared lysates were incubated with the M2 anti-Flag antibody-conjugated beads at 4 °C for 2 h and eluted with 200 μ g ml^{−1} 3× Flag peptide in buffer A (25 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10% glycerol) for 1 h. Recombinant RPA complex was purified from *E. coli* as previously described³⁰. hnRNPA1 pET9d plasmid (a gift from A. Krainer) was transformed into *E. coli* and expression was induced with IPTG (0.4 mM) for 3 h at 37 °C. The cells were then collected and lysed in binding buffer (10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 μ g ml^{−1} BSA, 10% glycerol, 0.05% NP-40). Lysates were sonicated, cleared by centrifugation (10,000g for 10 min) and incubated with ssTEL- (50 μ M) conjugated M280 beads (100 μ l) for 30 min at room temperature. The ssTEL and associated protein were captured by magnets, washed in binding buffer and eluted with 1 M NaCl for 10 min at 4 °C. The eluted protein was then diluted in binding buffer without salt to bring the final NaCl concentration down to 100 mM. *E. coli* single-stranded binding protein (SSB) was purchased from Promega.

Gel-shift assay. The 18-nucleotide telomeric ssDNA probe [(TTAGGG)₃] was radiolabelled with γ -³²P using T4 kinase and purified over a G25 column. The labelled ssDNA was incubated with purified RPA or POT1–TPP1 in binding buffer (10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 μ g ml^{−1} BSA, 10% glycerol, 0.05% NP-40) for 30 min at room temperature. The resulting protein–DNA complexes were separated by gel electrophoresis using 0.8% agarose at 140 V for 1.5 h and bands were visualized by autoradiography.

DNA–protein binding assay using biotinylated ssDNA. Biotinylated ssTEL [(TTAGGG)₈] or ssMUT [(TTTGCG)₈] were attached to streptavidin-coated magnetic beads in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl at room temperature for 30 min. To analyse the bindings of purified RPA, POT1–TPP1 and POT1 to ssDNA, biotinylated ssDNA (1 pmol) was incubated with various amounts of purified protein in 500 μ l of binding buffer. To analyse the binding of RPA and Flag–POT1 to ssDNA in extracts, biotinylated ssDNA (10 pmol) and various amounts of extracts were added to 500 μ l of binding buffer. After incubation for 30 min, the protein–DNA complexes were retrieved with a magnet and washed three times with binding buffer. In the experiments using RPA or POT1 pre-coated ssDNA, biotinylated ssDNA (1 pmol) was first incubated with purified protein (3.8 pmol) for 30 min at room temperature. The ssDNA pre-coated with RPA or POT1 was retrieved with a magnet and subsequently mixed with increasing concentrations of WCE, nuclear extract or cytoplasmic extract for 30 min at room temperature. For nuclear extract inhibited by addition of TERRA or its derivatives ((UUAGGG)₃, (CCCAUU)₃ and (UUGGCG)₃), extracts were incubated with 1, 2, 5 or 10 pmol RNA for 30 min at 4 °C.

For hnRNPA1 binding, RPA-coated ssTEL or ssMUT (0.8 nM), or Flag–POT1-coated ssTEL (0.8 nM), were incubated with increasing concentrations of hnRNPA1 purified from *E. coli* (2.4, 4.8, 7.2 nM) and the proteins remaining on ssTEL were analysed by western blot. For TERRA inhibition, hnRNPA1 was pre-incubated with increasing concentrations of TERRA (2, 4, 10, 20 nM), or control RNA (UUGGCG)₃. hnRNPA1 was then incubated with RPA coated ssTEL (0.8 nM). Similarly, to demonstrate that TERRA promotes the dissociation of hnRNPA1 from ssTEL, the ssTEL (0.8 nM) was pre-coated with hnRNPA1 (2.4 nM) and subsequently incubated with increasing concentrations of TERRA (2, 20, 200, 2,000 nM). To demonstrate that TERRA enhances POT1 binding, ssTEL (0.8 nM) was pre-coated with hnRNPA1 (2.4 nM) then incubated with both POT1 (2.4 nM) and increasing concentrations of TERRA (2, 20, 200 nM). In all reactions, the proteins remaining on DNA were analysed by western blot.

Cell synchronization. To follow the progression of cells from S to G2 (Fig. 4b, e), HeLa cells were synchronized with 2 mM thymidine for 16 h, washed three times with PBS and once with thymidine-free medium, and released into thymidine-free medium. To enrich HeLa cells in S phase of the cell cycle (Fig. 4a), cells were either collected after treatment for 16 h with 2 mM thymidine (early S), or collected 4 h after thymidine release (late S). To enrich cells in G1 and M phases, cells were either collected after treatment for 16 h with 0.1 μ g ml^{−1} nocodazole (M), or collected 4 h after nocodazole release (G1).

Extract preparation. WCEs were either generated with the NETN buffer as described in the protein purification section, or with the binding buffer used in the DNA binding assays. Nuclear extract and cytoplasmic extract were generated as previously described³. To treat extracts with TERRA or its derivative RNA, RNA was added to WCE or nuclear extract in increasing concentrations (1, 2, 5, 10 pmol) and incubated for 30 min on ice.

Capture of RPA-displacing activity from extracts. To capture the RPA-displacing activity from extracts, RPA-coated ssTEL was incubated with nuclear extract for 30 min at room temperature. The beads were collected, washed three times in binding buffer, and eluted using the binding buffer with 1 M NaCl for 10 min on ice. The eluted material was collected, and diluted with the binding buffer without NaCl to reach a final NaCl concentration of 100 mM. The elution was incubated on ice for 1 h then added to RPA-coated ssDNA and incubated for 30 min at room temperature. For TERRA inhibition, either TERRA (UUAGGG)₃ or its derivative (CCCAUU)₃ were incubated with the eluted proteins before their addition to RPA-coated ssDNA. The proteins remaining bound to DNA were analysed by western blot.

Identification of the RPA-displacing factors from extracts. Biotinylated ssTEL or ssMUT (20 pmol) was attached to streptavidin-coated beads and coated with recombinant RPA. The RPA-coated ssTEL or ssMUT was incubated with 65 μ g of nuclear extract in 500 μ l of binding buffer. Beads with no DNA attached were used as a negative control. After 30 min of incubation, the beads were retrieved and washed three times with binding buffer containing 300 mM NaCl. The proteins associated with the RPA activity were eluted by binding buffer containing 600 mM NaCl for 10 min on ice. The eluted proteins from ssTEL, ssMUT and naked beads were separated by SDS–PAGE. After the gel was silver-stained, the two ~30-kDa bands specifically captured by RPA–ssTEL were excised and analysed by mass spectrometry.

Immunofluorescence analysis. HeLa cells were seeded onto coverslips and cultured overnight. The adhered cells were transfected with POT1 siRNA using oligofectamine (Invitrogen), or with hnRNPA1 siRNA using Lipofectamine RNAi Max (Invitrogen) and cultured for another 48 h. Synchronized cells were treated after 24 h with 2 mM thymidine for 16 h, washed and released, and processed at the indicated time points. Cells were extracted with 0.25% Triton, fixed in 3% paraformaldehyde and further permeabilized with 0.5% Triton. Cells were subsequently incubated with the primary antibodies (diluted in PBS containing 3% BSA and 0.05% Tween 20) for 1 h at 37 °C in a humidified chamber. After extensive washing with PBS, cells were incubated with secondary antibodies for 45 min at room temperature, and washed again with PBS. After incubation for 5 min with DAPI, cells were mounted on slides with Vectashield. Slides were analysed using a Nikon H600L fluorescence microscope.

Combined immunofluorescence–RNA fluorescence in situ hybridization. Cells were grown on coverslips and collected at different time points 17 h after release of single thymidine block. Cells were washed twice with cold PBS for 5 min and treated with cyto buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 10 mM PIPES pH 7, 0.1% Triton X-100, 200 mM vanadyl ribonucleoside complex) for 7 min at 4 °C. Cells were rinsed briefly, fixed with 4% paraformaldehyde in PBS (USB 19943) for 10 min at room temperature. Cells were then washed three times with PBS for 5 min each and permeabilized with 0.5% NP40 in PBS for 10 min. Cells were washed twice with PBS for 5 min each and incubated with blocking solution (0.2% fish gelatin and 0.5% BSA) for 1 h. Cells were then incubated with human TRF2 antibody (clone 4A794 Upstate) at 1:2,000 and diluted in blocking solution for 2 h. After washing three times with PBST (PBS containing 0.1% Triton) for 10 min each, the cells were then incubated with secondary antibody Alexa 488 (Invitrogen A11001) at 1:2,000 dilution in blocking solution for 1 h. Cells were washed three times with PBST for 10 min each and were fixed with 4% paraformaldehyde in PBS for 10 min at room temperature. Cells were rinsed briefly with PBS then incubated with hybridization mix (10 nM PNA–TAMRA–(CCCTAA) probe, 50% formamide, 2× SSC, 2 mg ml^{−1} BSA, 10% dextran sulphate, 10 mM vanadyl ribonucleoside complex) for 18 h in a humidified chamber at 39 °C. Cells were washed with 2× SSC in 50% formamide three times at 39 °C for 5 min each, three times in 2× SSC at 39 °C for 5 min each, and finally once in 2× SSC at room temperature for 10 min. Coverslips were then mounted on glass microscope slides with Vectashield mounting medium containing DAPI (H-1200). For RNaseA treatment, coverslips were incubated with 200 μ g ml^{−1} RNase A for

30 min at 37 °C before hybridization. Images were captured with an Endore cooled CCD (charge-coupled device) camera on a Nikon eclipse 80i microscope and the images processed with NIS-Element BR 3.10 software.

Chromatin immunoprecipitation. RPA chromatin immunoprecipitation and the analysis of telomere association were performed as previously described⁷. Cells were transfected twice with hnRNPA1 siRNA (hnRNPA1-1) and synchronized with thymidine for 15 h. The two RPA2 antibodies used were from Abcam and Thermo.

Antibodies and siRNA. The RPA pS33 antibody was from Bethyl. The monoclonal antibody to RPA2 was from Neomarkers. The anti-FLAG M2 antibody was from Sigma. The Chk1 antibody and Cyclin A antibody were from Santa Cruz, and

the phospho-Chk1 Ser345 antibody was from Cell Signaling. The TRF2 antibody was from Bethyl. The phospho-H3 Ser10 antibody was from Millipore. The H4 antibody was from Active Motif. The hnRNPA1 antibody was from Cell Signaling. The POT1 siRNA used in Fig. 4e and Supplementary Fig. 10 was the SMARTPOOL from Dharmacon. The hnRNPA1 siRNAs used in Fig. 4c, d and Supplementary Figs 8 and 9 were CAACUUCGGUC-GUGGAGGA and UCCACGACCACCACCAAAG.

30. Henricksen, L. A. & Wold, M. S. Replication protein A mutants lacking phosphorylation sites for p34cdc2 kinase support DNA replication. *J. Biol. Chem.* **269**, 24203–24208 (1994).

Caspase signalling controls microglia activation and neurotoxicity

Miguel A. Burguillos^{1,2,3}, Tomas Deierborg³, Edel Kavanagh¹, Annette Persson⁴, Nabil Hajji^{1†}, Albert Garcia-Quintanilla⁵, Josefina Cano⁵, Patrik Brundin³, Elisabet Englund⁴, Jose L. Venero² & Bertrand Joseph¹

Activation of microglia and inflammation-mediated neurotoxicity are suggested to play a decisive role in the pathogenesis of several neurodegenerative disorders. Activated microglia release pro-inflammatory factors that may be neurotoxic. Here we show that the orderly activation of caspase-8 and caspase-3/7, known executioners of apoptotic cell death, regulate microglia activation through a protein kinase C (PKC)- δ -dependent pathway. We find that stimulation of microglia with various inflammogens activates caspase-8 and caspase-3/7 in microglia without triggering cell death *in vitro* and *in vivo*. Knockdown or chemical inhibition of each of these caspases hindered microglia activation and consequently reduced neurotoxicity. We observe that these caspases are activated in microglia in the ventral mesencephalon of Parkinson's disease (PD) and the frontal cortex of individuals with Alzheimer's disease (AD). Taken together, we show that caspase-8 and caspase-3/7 are involved in regulating microglia activation. We conclude that inhibition of these caspases could be neuroprotective by targeting the microglia rather than the neurons themselves.

Numerous *in vivo* clinical imaging and neuropathology studies suggest that activated microglia, the resident immune cells of the central nervous system, play prominent roles in the pathogenesis of neurodegenerative disorders, including PD, multiple sclerosis and AD^{1,2}. Microglia are necessary for normal brain function; however, uncontrolled and over-activated microglia can trigger neurotoxicity. They are a prominent source of pro-inflammatory factors and oxidative stress such as tumour-necrosis factor (TNF)- α , nitric oxide and interleukin (IL)-1 β , which are neurotoxic^{2,3}.

Toll-like receptors (TLRs) are a family of pattern-recognition receptors in the innate immune system. Exogenous and endogenous TLR ligands activate microglia⁴. Intracerebral delivery of lipopolysaccharide (LPS), the major component of Gram-negative bacterial walls and a ligand for TLR4, leads *in vivo* to microglia activation and neuronal injury, and is used as model for brain inflammation^{4,5}. Synergistic effects between interferon- γ (IFN- γ) and several TLR ligands (including TLR4) have been suggested, suggesting crosstalk between these pro-inflammatory receptor signalling pathways⁶. Furthermore, IFN- γ receptor-deficient mice are less susceptible to LPS-induced endotoxic shock than control mice⁷. Finally, TLR4 has been implicated in AD pathophysiology in several contexts. Thus, the upregulation of cytokines is TLR4 dependent in an AD mouse model⁸; certain TLR4 single nucleotide polymorphisms are associated with increased risk for AD⁹; the levels of TLR4 messenger RNA (mRNA) are upregulated in APP transgenic mice¹⁰; and increased TLR4 expression is associated with amyloid plaque deposition in AD brain tissue¹⁰.

Caspases, a family of cysteinyl-aspartate-specific proteases, are executioners of apoptotic cell death and their activation is considered a commitment to cell death^{11,12}. Certain caspases, for example caspase-1, also play a pivotal role in immune-mediated inflammation. In this situation, caspase activation is associated with the maturation of pro-inflammatory cytokines, such as IL-1 β , IL-18, IL-33, and not with apoptosis¹³. Inhibition of caspase activation protects against

neuronal loss in several animal models of brain diseases involving activated microglia, including hypoxic ischaemia/stroke, acute bacterial meningitis, brain trauma and 6-hydroxydopamine and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-lesioned parkinsonism models^{2,14–17}. Currently, it is unclear whether inhibition of caspase activation specifically in microglia contributes to the neuroprotective effects of caspase inhibitors. We have now discovered that microglial activation in cell and animal models of inflammation involves caspases and that inhibition of the cascade in microglia prevents neurodegeneration. Furthermore, we demonstrate that caspase activation occurs in microglia in the brains of individuals with PD and AD, and thereby we validate the observations we made in relevant cell and animal models.

Results

Caspase-3/7 control microglia activation

We stimulated BV2 cells with LPS to investigate the molecular pathways involved in microglia activation (Fig. 1). LPS treatment induced caspase-3 cleavage and D(OMe)E(OMe)VD(OMe)-ase (DEVD-ase) activity after only 4 h in BV2 microglia cells in a time- and dose-dependent manner (Fig. 1a and Supplementary Figs 2a, b, e, f and 5d, e). Increased DEVD-ase activity was also observed upon treatment with other pro-inflammogens such as lipoteichoic acid (LTA, TLR2 agonist), PamC3sk4 (synthetic lipopeptide TLR1/2 agonist) and interferon-gamma (IFN- γ) (Supplementary Fig. 3a). This activity primarily reflects caspase-3/7 activities. Both caspase 3 and 7 are known as major apoptosis executioners. Despite the increase in DEVD-ase activity, we did not observe major microglia cell death within the first 24 h after initiating LPS treatment (Fig. 1c and Supplementary Figs 2c and 4a–c, e, f) nor with LTA, PamC3sk4 and IFN- γ treatments (Supplementary Fig. 3b, c). Thus, the LPS-induced increase in caspase 3 and 7 activity did not result in major cell death and the little cell death occurring at 48 h was not prevented using the caspase-3/7 inhibitor Z-D(OMe)E(OMe)VD(OMe)-FMK

¹Department of Oncology-Pathology, Cancer Centrum Karolinska, Karolinska Institutet, 171 76, Stockholm, Sweden. ²Departamento de Bioquímica y Biología Molecular, Facultad de Farmacia, Universidad de Sevilla, and Instituto de Biomedicina de Sevilla, 41012 Sevilla, Spain. ³Neuronal Survival Unit, Wallenberg Neuroscience Center, Department of Experimental Medical Science, 221 84 Lund, Sweden. ⁴Department of Pathology, Division of Neuropathology, Lund University Hospital, 221 85 Lund, Sweden. ⁵Servicio de Biología, Centro de Investigación, Tecnología e Innovación, Universidad de Sevilla (CITIUS), 41012 Sevilla, Spain. [†]Present address: Department of Experimental Medicine and Toxicology, Division of Investigative Science, Imperial College London, UK.

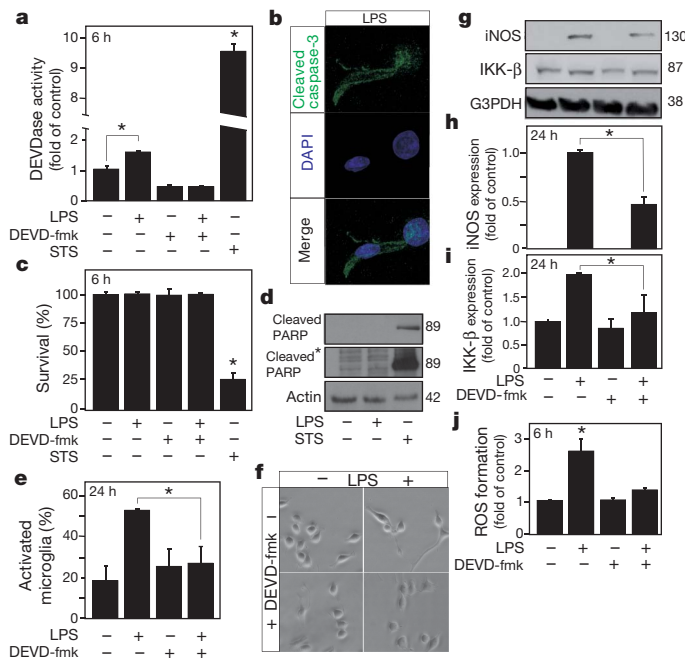


Figure 1 | LPS-induced DEVDase activity regulates microglia activation but not cell death. LPS treatment induces DEVDase activity (a) and processing of caspase-3 (b), which are not associated with cell death, as illustrated by cell survival quantification (c), and absence of PARP cleavage (d) in cultured BV2 microglia cells. Caspase-3/7 inhibition by DEVD-fmk (a) prevents LPS-induced morphological microglia activation (e, f) and induction of typical inflammation-related molecules like iNOS (g, h) and IKK- β (g, i) and ROS formation (j). STS is used as an apoptosis inducer. Data are expressed as mean \pm s.e.m. ($n = 3$). * $P < 0.05$.

(DEVD-fmk) (Supplementary Figs 2d and 4d, g). Compared with LPS treatment, exposure of BV2 microglia cells to a death stimulus such as staurosporine (STS) led to a significantly greater caspase-3 cleavage and induction of DEVDase activity (Fig. 1a and Supplementary Fig. 5d, e). After LPS treatment, we found cleaved caspase-3 to be located primarily close to the plasma membrane and not present in the nucleus (Fig. 1b, Supplementary Fig. 5a–d and Supplementary Movies 1 and 2). Furthermore, we did not observe cleavage of the caspase-3/7 nuclear substrate poly(ADP-ribose) polymerase (PARP-1) in response to LPS (Fig. 1d) even after long exposure of the membrane (asterisk in Fig. 1d). In addition, whereas exposure of BV2 cells to STS promoted Bid processing and loss of mitochondrial transmembrane potential, these two events associated with the mitochondrial cell death pathway were found to be unaffected upon LPS treatment (Supplementary Fig. 10a–c). When we inhibited DEVDase activity in BV2 cells by exposing them to the cell-permeable and irreversible caspase inhibitor DEVD-fmk, LPS treatment failed to activate the microglia. Thus, the cells did not exhibit morphological changes associated with microglia activation (Fig. 1e–f) and did not show features of activated microglia, such as I κ B kinase complex β (IKK- β), inducible nitric oxide synthases (iNOS) and reactive oxygen species (ROS) formation (Fig. 1g–j). Co-treatment with DEVD-fmk also prevented the LTA, PamC3sk4 and IFN- γ -induced iNOS expression and ROS formation (Supplementary Fig. 6a–e).

Caspase-3 and -7 account for cellular DEVDase activity. We therefore decided to assess their respective roles in microglia activation by selectively knocking down endogenous caspase-3 (Supplementary Fig. 7a, b) or caspase-7 (Supplementary Fig. 7c, d) using a pool of small interfering RNAs (siRNAs). First, we confirmed that the silencing of these proteases effectively decreased DEVDase activity (Supplementary Fig. 7e). Then, when we transfected BV2 microglia cells with siRNA targeting specifically either one of the two caspases, LPS-treatment did not induce iNOS, IKK- β expression, ROS formation and production of certain cytokine production (IL-1 β , TNF- α and

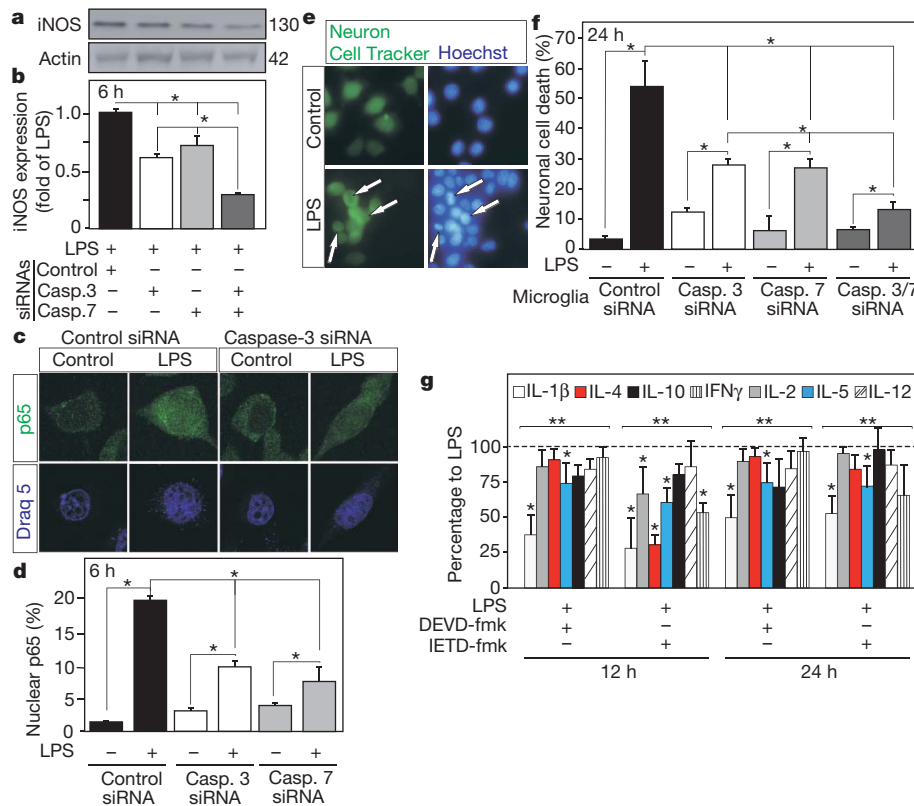
murine keratinocyte chemoattractant) as effectively (Fig. 2a, b and Supplementary Fig. 7f–i). We found that silencing both caspases simultaneously reduced IKK- β expression even further (Supplementary Fig. 7f). This indicates that DEVDase activity per se regulates microglia activation (Fig. 2 and Supplementary Fig. 7). Similar responses to caspase-3 and -7 knockdowns were observed with LTA, PamC3sk4 and IFN- γ treatments (Supplementary Fig. 8a–i). Activation and nuclear translocation of nuclear factor κ B (NF- κ B) is a key step in LPS-induced microglia activation¹⁸. NF- κ B is sequestered in the cytoplasm by the I κ B family of inhibitory proteins that mask the nuclear localization signal of NF- κ B. IKK- β can phosphorylate I κ B, and thereby target it for degradation through the ubiquitin proteasome pathway. As a consequence, functional NF- κ B molecules then become free to enter the nucleus. We detected less nuclear NF- κ B p65 subunits in cells subjected to knockdown of caspase-3 or caspase-7 before LPS treatment (Fig. 2c, d), indicating that reduced caspase activation led to less nuclear translocation of NF- κ B. Finally, using microglia cells co-cultured with dopaminergic neurons, we examined whether the inhibition of the IKK/NF- κ B pathways by selective knockdown of caspase-3 or/and caspase-7 was associated with loss of microglia neurotoxicity. In agreement with earlier studies¹⁹, LPS treatment activated microglia and caused dopaminergic neurons to die. Importantly, we found that reducing the LPS-induced microglia activation by caspase knockdown meant that fewer co-cultured dopaminergic neurons died (Fig. 2e, f). Also, we checked the level of several cytokines in primary microglial cell cultures at 12 and 24 h, and observed a decrease of IL-1 β and IL-5 with both DEVD-fmk and IETD-fmk, and of IL-2 and IFN- γ at 12 h when we used IETD-fmk (Fig. 2g).

Caspase-8 directs caspase-3/7 activation

We then examined how caspase-3/7 are activated in LPS-treated microglia. LPS treatment has been reported to promote caspase-1 activation²⁰. Caspase 1 is a key component of the inflammasome required for the processing and maturation of pro-inflammatory cytokines. It plays a pivotal role during LPS-induced inflammation. Therefore caspase-1 null mice and mice expressing a dominant-negative mutant caspase-1 gene exhibit reduced LPS-induced inflammation^{21,22}. Consequently, we examined if caspase-1 acts upstream of caspase-3/7 in LPS-induced microglia activation. We found that the caspase-1 inhibitor YVAD-fmk did not reduce LPS-induced activation of microglia, as assessed by DEVDase activity (Fig. 3a).

Caspase-8 is believed to be at the apex of the death receptor-mediated apoptosis pathway and can activate caspase-3/7 (refs 23–25). We found that LPS induced caspase-8 activity (IETDase) within 6 h of being added to BV2 microglia cultures (Fig. 3b, c). Consistent with this result, we observed that the caspase-8 inhibitor IETD-fmk or caspase-8 knockdown using siRNA prevented LPS-induced DEVDase activity (Fig. 3a, d). The IETD-fmk treatment or caspase-8 silencing only prevented the LPS-induced increase in DEVDase activity, whereas the caspase inhibitor DEVD-fmk reduced the DEVDase activity even further (Fig. 3a, c). Knockdown or chemical inhibition of caspase-8 were also associated with a reduction of iNOS expression and ROS formation upon treatment with all tested pro-inflammogens (that is, LPS, LTA, PamC3sk4 and IFN- γ) (Fig. 3e, f and Supplementary Figs 7i, 8a–i and 9a, b). Other known potential caspase-8 substrates, namely Bid, HDAC7 and RIP1, were found not to be processed upon LPS treatment (Supplementary Fig. 10 b–f). Taken together, this indicates that the LPS-induced increase in DEVDase activity is dependent on caspase-8 activation. Furthermore, we found that the caspase-8- and caspase-1-initiated pathways had additive effects in the regulation of LPS-induced iNOS expression (Supplementary Fig. 9c, d). This suggests that both of these caspase-regulated pathways contribute to LPS-induced inflammation. We also examined if the observed DEVDase activity and that of caspase 6, the remaining executioner caspase, were correlated. However, we were unable to

Figure 2 | Knockdown of caspase-3 or caspase-7 decreases microglia activation in response to LPS. siRNAs targeting caspase-3 and/or caspase-7 prevent LPS-induced increase of iNOS expression (a, b) in cultured BV2 cells. In addition, caspase-3 or caspase-7 silencing prevents LPS-induced activation of NF- κ B as seen by reduction of p65 nuclear translocation (c, d). Under co-cultured conditions, LPS treatment induces neuronal cell death and caspase-3/7 knockdown in microglia cells prevents neuronal cell death (e, f). DEVD-fmk or IETD-fmk (caspase-8 inhibitor) treatment modulates cytokine expression at 12 and 24 h in primary microglia cell culture (g). Data are expressed as mean \pm s.e.m. ($n = 3$) and \pm s.d. ($n = 4$) in g. * $P < 0.05$. In g, ** denotes statistically significant difference between treatments ($P < 0.001$) concerning time and treatment.



detect any VEID-ase activity (related to caspase-6) in cells upon LPS treatment. Pretreatment with VEID-fmk, a caspase 6 inhibitor, did not affect LPS-induced DEVD-ase activity, indicating that this protease does not play an essential role in the activation process (Supplementary Fig. 11a, b).

The caspase signalling depends on TLR4

We then examined further the link between TLR4 ligation and activation of caspase-8. Selective knockdown of TLR4 (Supplementary Fig. 12a) was associated with a reduced caspase-8 activation and consequent caspase-3 activation, providing evidence for a direct activation of caspase-8 by TLR4 ligation (Supplementary Fig. 12b, c). Apoptosis after TLR2 activation has been reported to be associated to the formation of a myeloid differentiation factor 88 (MyD88)/Fas-associated

death domain protein (FADD)/Caspase-8 complex²⁶. To assess the potential role of this complex in TLR4 ligation-induced microglia activation, we knocked down MyD88 by siRNAs (Supplementary Fig. 12d). We observed that downregulation of MyD88 did not affect LPS-induced activation of caspase-3 and caspase-8 (Supplementary Fig. 12e, f), suggesting that TLR4 ligation-induced microglia activation did not act through recruitment of the MyD88/FADD/caspase-8, as is the case for TLR2 ligation-induced apoptosis^{26,27}.

Microglia activation can be regulated by autocrine signalling of TNF- α by the TNF receptor 1 (TNFR1) secreted by LPS-stimulated microglia²⁸. To examine whether we were observing a secondary effect, induced by autocrine signalling of TNF- α excreted by LPS-stimulated microglia cells, we investigated the effect of neutralizing TNF- α receptor antibodies on LPS-induced caspase-3 and caspase-8 activities²⁸ (Supplementary Fig. 13a-d). Selectively blocking the TNF- α receptor resulted in decreased LPS-induced iNOS expression at 24 h, thus confirming the TNF- α positive feedback loop exerted on microglia cells (Supplementary Fig. 13e). However, it failed to affect the LPS-induced caspase-3 and caspase-8 activities (Supplementary Fig. 13a-d). Thus, this experiment demonstrates that the caspase-signalling pathway is directly activated, independently of TNF- α receptor stimulation, upon LPS treatment.

Cleaved PKC- δ mediates effect of caspases

We then explored how activated caspase-3/7 interacts with the IKK/NF- κ B pathway during LPS-induced microglia activation. PKC- δ has been reported to regulate NF- κ B activation through the IKK complexes and phosphorylation of the NF- κ B inhibitor I κ B^{29,30}. Interestingly, PKC- δ can be cleaved by caspases to generate a 40 kDa catalytically active fragment³¹. Because the levels of this protein are quite low in this cell line, we decided to overexpress PKC- δ to detect the cleaved form (Supplementary Fig. 14a). We found that LPS treatment of BV2 microglia cells promoted cleavage of PKC- δ into its 40-kDa active fragment (Fig. 4a, b). We found that the caspase inhibitor DEVD-fmk or selective siRNA knockdown of caspase-3 or caspase-7 reduced LPS-induced PKC- δ activation. Therefore, in this context, PKC- δ activation is dependent on DEVD-ase activity (Fig. 4c, d). We then obtained further

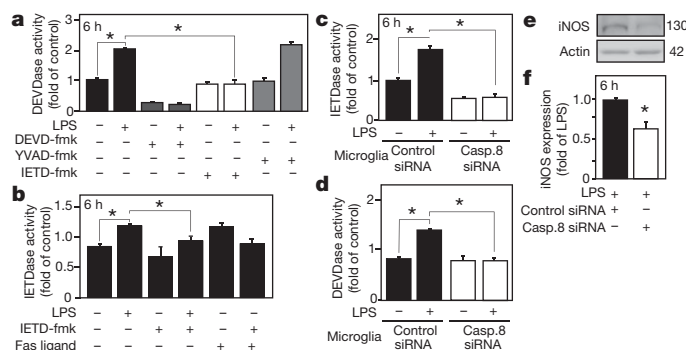


Figure 3 | Caspase-8 activity controls LPS-induced caspase-3/7 activation. Caspase-8 inhibition by IETD-fmk, but not caspase-1 inhibition by YVAD-fmk, prevents LPS-induced DEVDase activity (a). LPS treatment induces caspase-8 (IETDase) activity, which is inhibited by IETD-fmk (b) or specific siRNA targeting caspase-8 (c) in BV2 cells. Treatment with Fas ligand was used as positive control for IETDase activity induction (b). Caspase-8 knockdown using siRNA prevents LPS-induced DEVDase activity (d). siRNA knockdown of caspase-8 prevents LPS-induced expression of iNOS (e, f). Data are expressed as mean \pm s.e.m. ($n = 3$). * $P < 0.05$.

evidence supporting that PKC- δ activation is important in microglia activation. First, we found that the PKC- δ inhibitor rottlerin inhibits LPS-induced iNOS expression in microglia (Fig. 4e, f). Second, if we overexpressed in microglia cells a caspase-uncleavable mutant of PKC- δ , we observed a decrease of LPS-induced iNOS expression (Fig. 4g, h). By contrast, when we overexpressed PKC- δ , we further enhanced the activating effect of LPS on microglia (Supplementary Fig. 14b). Also, we observed a marked decrease of all cytokines in primary microglial cell cultures after being challenged with rottlerin (Fig. 4i). Taken together, our experiments demonstrate that the caspase-3/7-dependent activation of microglia by LPS is mediated through PKC- δ .

Role of the caspase signalling *in vivo*

To examine the physiological relevance of these findings, we performed *in vivo* experiments and injected LPS into the rat substantia nigra. The substantia nigra is known to exhibit a strong inflammatory response upon LPS challenge⁴. At 24 h after injection (Fig. 5a–c), we observed a strong induction of caspase-8 and activation of caspase-3, which was mostly confined to reactive microglia in the mesencephalon on the injected side (Fig. 5b, c). By contrast, microglia were quiescent in the contralateral control midbrain (Fig. 5a). To study if caspase-3/7 activation is important for microglia activation also *in vivo*, we co-injected DEVD-fmk with LPS. Twenty-four hours later, we found that caspase-3/7 inhibition prevented LPS-induced microglia activation (Supplementary Fig. 15a, b). At the molecular level, we observed that the DEVD-fmk treatment mitigated LPS-induced expression of cytokines and pro-inflammatory molecules including iNOS, TNF- α and IL-1 β (Fig. 5d). As a final step, we examined if our earlier *in vitro* findings about the role of PKC- δ in caspase-mediated microglia activation are relevant *in vivo*. We injected LPS intranigally with IETD-fmk, DEVD-fmk or rottlerin. Twenty-four hours later, we dissected the substantia nigra and measured the iNOS protein level. In agreement with our *in vitro* data, we found that inhibition of caspase-8, caspase-3/7 or PKC- δ activities significantly decreased LPS-induced iNOS expression (Fig. 5e, f).

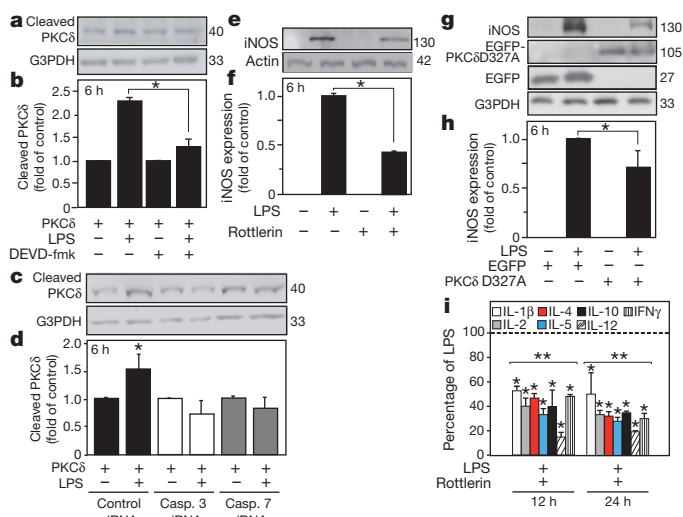


Figure 4 | Caspase-3/7 regulates microglia activation through the PKC- δ pathway. Caspase-3/7 inhibition by DEVD-fmk (**a**, **b**) or by siRNA knockdown (**c**, **d**) prevents LPS-induced PKC- δ cleavage/activation in BV2 microglia cells. PKC- δ inhibition by rottlerin mostly abolishes LPS-induced activation of iNOS (**e**, **f**). Overexpression of PKC- δ D327A, a caspase uncleavable PKC- δ mutant, has an inhibitory effect on LPS-induced iNOS expression (**g**, **h**). Rottlerin treatment reduces LPS-induced cytokine expression at 12 and 24 h in primary microglia cell culture (**i**). Data are expressed as mean \pm s.e.m. ($n = 3$) and \pm s.d. ($n = 4$) in **i**. * $P < 0.05$. In **i**, ** denotes statistically significant difference between treatments $P < 0.001$, concerning time and treatment.

To substantiate further our observations in the rat LPS model, we monitored the activation of microglia in the MPTP-lesion mouse model of PD. Consistent with earlier work, mice injected systemically with MPTP showed a strong microglial activation and a robust neurotoxicity in ventral mesencephalic dopamine neurons³². Intranigral vehicle injections in sham animals (saline containing 1% DMSO) greatly increased the density of reactive microglia, as evidenced by Iba1 immunohistochemistry (Supplementary Fig. 15c, d). Consequently, we quantified the numbers of reactive and resting microglia in Iba1-immunostained sections covering injection and non-injection sites for each experimental condition (Fig. 5g, h and Supplementary Fig. 15c, d). As expected, the density of reactive microglia at the non-injection site was highest in the MPTP group and dramatically lower in the sham (no MPTP) group. Interestingly, caspase 8 inhibition significantly prevented the MPTP-induced microglia activation and the MPTP-induced reduction in the density of resting microglia (Fig. 5g, h). Similar results of IETD-fmk were observed at the injection site (Supplementary Fig. 14c, d). As expected, MPTP injections severely reduced the integrity of nigro-striatal dopaminergic terminals (Supplementary Fig. 15e–g). Intranigral IETD-fmk injection induced a modest, but significant, protection against MPTP-induced toxicity as demonstrated by densitometric analysis of surviving striatal dopaminergic terminals and stereological cell counts in the substantia nigra (Supplementary Fig. 15e–g).

Caspase activation in human brains

PD and AD are known to be associated with neuroinflammation and the presence of activated microglia^{2,33}. We investigated whether caspase-3 and caspase-8 are activated in microglia of individuals with PD and AD. We analysed expression of cleaved caspase-3, cleaved caspase-8 and the microglia marker CD68 in post-mortem brain samples from individuals with PD and AD (clinically as well as neuropathologically diagnosed) and age- and gender-matched healthy control brains. We detected significant cytoplasmic expression of both active caspase-3 and active caspase-8 in the PD ventral mesencephalon and in the AD frontal cortex, compared with controls (Fig. 6a–d and Supplementary Figs 16–19). The activated caspases and the microglial marker CD68 were largely co-localized, indicating that caspase-8 and caspase-3 are activated mainly in microglia in PD and AD (Fig. 6a–d and Supplementary Fig. 20a, b).

Discussion

In this study, we uncover a completely novel role for caspase-8 and caspase-3/7 in the control of microglia and brain inflammation. We show that stimulation of microglia with the pro-inflammatory stimuli LPS, LTA, PamC3sk4 and IFN- γ triggers caspase-3/7 activation, without causing cell death. Caspases are proteases essential for apoptosis and inflammation. Caspase-1 is already known as the prototypical inflammatory caspase, required for the maturation of pro-inflammatory cytokines. By comparison, the initiator caspase-8 and effector pro-apoptotic caspase-3 and caspase-7 are considered crucial in the intracellular death machinery. Unexpectedly, we showed that caspase-3/7-dependent DEVDase activity controls LPS-induced microglia activation. We demonstrated that inhibition of the caspase-3/7 pathway effectively block microglia activation. For example, we observed fewer microglia with an activated phenotype in the presence of the caspase inhibitor DEVD-fmk and inhibition of the downstream IKK/NF- κ B pathways. We also reveal that caspase-3/7 activates the IKK/NF- κ B pathways through processing and activation of PKC- δ . Furthermore, we found that microglia exposed to the pro-inflammatory agent LPS failed to be toxic to neighbouring neurons when we inhibited caspase-3/7 chemically or by siRNA gene silencing. Importantly, we provide compelling evidence that active caspases 3 and 8 are expressed within reactive microglia in the ventral mesencephalon and frontal cortex of individuals with PD and AD, respectively.

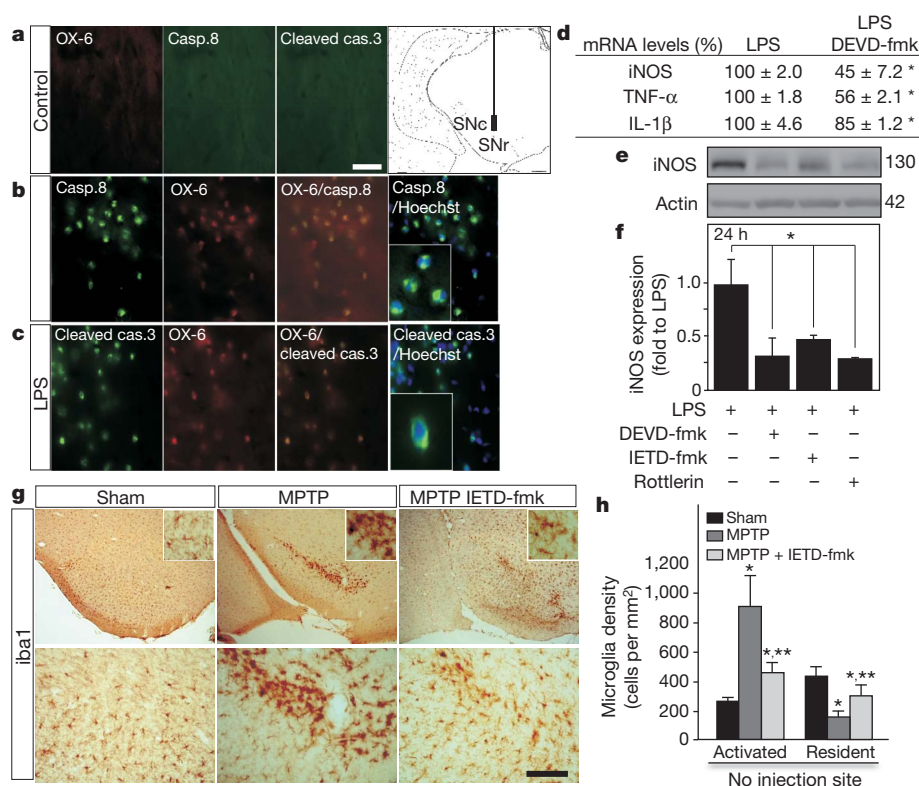


Figure 5 | *In vivo* inhibition of the caspase-dependent pathways prevents microglia activation. *In vivo* injection of LPS in rat substantia nigra (insert in **a**) induced activation of caspase-8 and caspase-3 in the ventral mesencephalon, which is mostly localized in OX-6-labelled reactive microglia (**b**, **c**) compared with control (**a**). There is a null coincidence between Hoechst-labelled nuclei and caspase8/cleaved caspase 3 immunoreactivity in OX-6-labelled microglia (**b**, **c**). *In vivo* DEVD-fmk significantly reduced intranigral LPS-induced microglia activation as evaluated by mRNA levels of proinflammatory molecules (**d**). *In vivo* inhibition of caspase-8, caspase-3/7 or PKC-δ, using IETD-fmk, DEVD-fmk and Rottlerin respectively, predominantly abolished the intranigral induction of iNOS upon LPS treatment (**e**, **f**). *In vivo* acute intraperitoneal injections of MPTP in mice highly increased the density of reactive microglia in substantia nigra (**g**, **h**). Intranigral IETD-fmk injections robustly prevented the MPTP-induced microglia activation and decrease of the resident microglia density at the non-injection site (**g**, **h**). Data are expressed as mean ± s.e.m. ($n = 3$). SNr, substantia nigra pars reticulata; SNc, Substantia nigra pars compacta. Scale bars: **a–c**, 25 μm; **d**, 500 μm for low magnification photographs; **j**, 400 and 100 μm for low- and high-magnification photographs. * $P < 0.05$; ** denotes statistically significant difference between MPTP in **h**, and MPTP+IETD-fmk treatments.

To summarize, we present new non-apoptotic functions for caspases-8, -3 and -7 and show that they can have a pivotal role in inflammation of the central nervous system (Supplementary Fig. 1). Brain inflammation is a typical feature of neurodegenerative diseases^{1,2,34} and is a prominent sequel of many acute forms of brain injury (for example, trauma, encephalitis and stroke)^{35,36}. Under certain circumstances, neuroinflammation is known to promote neuronal death³⁷. Accordingly, previous studies have shown that

anti-inflammatory treatment can reduce neurodegeneration. Our discovery that the caspases-8, -3 and -7 cascade can promote neuroinflammation through IKK/NF-κB and PKC-δ, with the development of nanocarriers that allow the inhibitor to cross the blood–brain barrier, opens up new molecular targets for anti-inflammatory drugs³⁸. Considering that these caspases also regulate apoptotic neuronal death, our results should revitalize interest in caspase inhibitors as potential therapeutic agents in disorders of the central nervous system^{39–41}.

METHODS SUMMARY

Microglial BV2 and dopaminergic MN9D murine cells were cultured as described^{42,43}. Experiments were performed in reduced 5% FCS media. Transfection of BV2 cells used Lipofectamine 2000 (Invitrogen). Non-targeting control, caspase-3, caspase-7 and caspase-8 siRNA were obtained from Dharmacon. Single and co-cultures were exposed to 1 μg ml⁻¹ LPS. Primary microglial cells were prepared from postnatal day 1–3 mouse brain using a previously described protocol⁴⁴. Cytokine content was assayed using the Mouse TH1/TH2 9-PlexTissue Culture Kit (Meso-scale Discovery). DEVD-fmk (20 μM), IETD-fmk (20 μM), YVAD-fmk (20 μM), rottlerin (2 μM) or TNF R1 mouse Ab (20 ng ml⁻¹) were added to the media 1 h before LPS treatment. Under approved protocols, male albino Wistar rats (230–250 g) were intranigally injected with 2 μg LPS alone or in combination with 0.75 nmol DEVD-fmk or 0.75 nmol IETD-fmk. Rottlerin (20 mg kg⁻¹) was administered intraperitoneally. Twenty-four hours after surgery, rats were killed and brains processed for analysis. Male C57BL mice were intranigally injected with vehicle or 0.75 nmol IETD-fmk. Twelve hours later, animals were treated with four injections of MPTP (16 mg kg⁻¹) at 2-h intervals. Four days after the last injection, animals were killed and brains processed for analysis. Paraffin-embedded archival tissue blocks from autopsy on three individuals with PD and three individuals with AD were obtained from the Department of Pathology, Lund University Hospital, Sweden. Age-matched controls (cardiac arrest victims, no brain disease) were analysed at the same time. Histological, immunological and fluorescence-activated cell sorting (FACS) analyses using antibodies listed in Supplementary Table 1 were performed using standard procedures^{42,45–47}. The subcellular localization of cleaved caspase-3, cleaved caspase-8 or NF-κB p65 protein was determined by confocal microscopy. Quantitative PCR with primers listed in Supplementary Table 2 were performed using SensiMixPlus SYBR (Quantace). MTT, caspase activity assay, apoptosis quantification and FACS analysis have been previously described⁴⁵. Statistical evaluations were performed by one- or two-way analysis of variance with Bonferroni post hoc tests and Kolmogorov–Smirnov test.

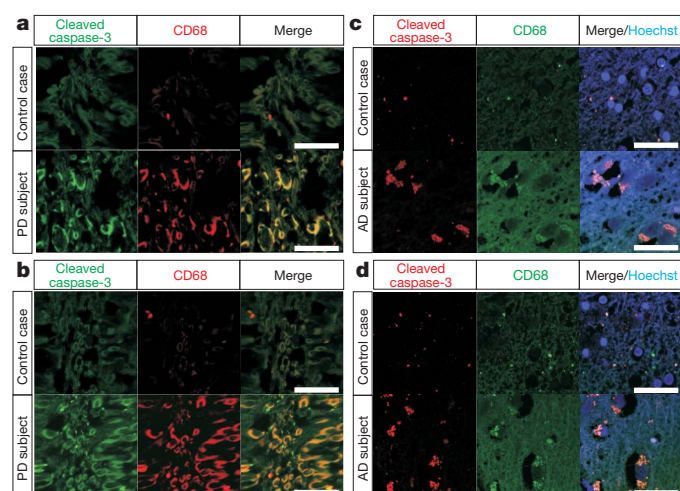


Figure 6 | Activation of caspase-3 and caspase-8 in microglia in brain from individuals with PD and AD. Double immunolabelling and confocal imaging analysis show significant activation of caspase-8 and caspase-3, mostly localized in CD68-labelled microglia, in the ventral mesencephalon from patients with PD (**a**, **b**) and in the frontal cortex of patients with AD (**c**, **d**) compared with age- and gender-matched healthy controls. Confocal images of immunostaining with cleaved caspase-3 (**a**, **c**) or cleaved caspase-8 (**b**, **d**), together with CD68 antibodies, are depicted. Merge from a control and an individual with AD illustrates the absence of cleaved caspase 3/caspase staining in Hoechst-labelled nuclei in reactive microglia (**c**, **d**). Scale bar, 10 μm.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 October; accepted 23 December 2010.

Published online 9 March 2011.

- Hanisch, U. K. & Kettenmann, H. Microglia: active sensor and versatile effector cells in the normal and pathologic brain. *Nature Neurosci.* **10**, 1387–1394 (2007).
- Block, M. L., Zecca, L. & Hong, J. S. Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. *Nature Rev. Neurosci.* **8**, 57–69 (2007).
- Chao, C. C., Hu, S., Molitor, T. W., Shaskan, E. G. & Peterson, P. K. Activated microglia mediate neuronal cell injury via a nitric oxide mechanism. *J. Immunol.* **149**, 2736–2741 (1992).
- Castano, A., Herrera, A. J., Cano, J. & Machado, A. Lipopolysaccharide intranigral injection induces inflammatory reaction and damage in nigrostriatal dopaminergic system. *J. Neurochem.* **70**, 1584–1592 (1998).
- Saijo, K. et al. A Nurr1/CoREST pathway in microglia and astrocytes protects dopaminergic neurons from inflammation-induced death. *Cell* **137**, 47–59 (2009).
- Zhao, J. et al. IRF-8/interferon (IFN) consensus sequence-binding protein is involved in Toll-like receptor (TLR) signaling and contributes to the cross-talk between TLR and IFN-gamma signaling pathways. *J. Biol. Chem.* **281**, 10073–10080 (2006).
- Car, B. D. et al. Interferon gamma receptor deficient mice are resistant to endotoxic shock. *J. Exp. Med.* **179**, 1437–1444 (1994).
- Jin, J. J., Kim, H. D., Maxwell, J. A., Li, L. & Fukuchi, K. Toll-like receptor 4-dependent upregulation of cytokines in a transgenic mouse model of Alzheimer's disease. *J. Neuroinflammation* **5**, 23 (2008).
- Balistreri, C. R. et al. Association between the polymorphisms of TLR4 and CD14 genes and Alzheimer's disease. *Curr. Pharm. Des.* **14**, 2672–2677 (2008).
- Walter, S. et al. Role of the toll-like receptor 4 in neuroinflammation in Alzheimer's disease. *Cell. Physiol. Biochem.* **20**, 947–956 (2007).
- Nicholson, D. W. et al. Identification and inhibition of the ICE/CED-3 protease necessary for mammalian apoptosis. *Nature* **376**, 37–43 (1995).
- Cohen, G. M. Caspases: the executioners of apoptosis. *Biochem. J.* **326**, 1–16 (1997).
- Keller, M., Ruegg, A., Werner, S. & Beer, H. D. Active caspase-1 is a regulator of unconventional protein secretion. *Cell* **132**, 818–831 (2008).
- Schulz, J. B. et al. Extended therapeutic window for caspase inhibition and synergy with MK-801 in the treatment of cerebral histotoxic hypoxia. *Cell Death Differ.* **5**, 847–857 (1998).
- Braun, J. S. et al. Neuroprotection by a caspase inhibitor in acute bacterial meningitis. *Nature Med.* **5**, 298–302 (1999).
- Cutillas, B., Espejo, M., Gil, J., Ferrer, I. & Ambrosio, S. Caspase inhibition protects nigral neurons against 6-OHDA-induced retrograde degeneration. *Neuroreport* **10**, 2605–2608 (1999).
- Depino, A. M. et al. Microglial activation with atypical proinflammatory cytokine expression in a rat model of Parkinson's disease. *Eur. J. Neurosci.* **18**, 2731–2742 (2003).
- Kawai, T. & Akira, S. Signaling to NF- κ B by Toll-like receptors. *Trends Mol. Med.* **13**, 460–469 (2007).
- Gibbons, H. M. & Dragunow, M. Microglia induce neural cell death via a proximity-dependent mechanism involving nitric oxide. *Brain Res.* **1084**, 1–15 (2006).
- Schumann, R. R. et al. Lipopolysaccharide activates caspase-1 (interleukin-1-converting enzyme) in cultured monocytic and endothelial cells. *Blood* **91**, 577–584 (1998).
- Li, P. et al. Mice deficient in IL-1 β -converting enzyme are defective in production of mature IL-1 β and resistant to endotoxic shock. *Cell* **80**, 401–411 (1995).
- Friedlander, R. M. et al. Expression of a dominant negative mutant of interleukin-1 β converting enzyme in transgenic mice prevents neuronal cell death induced by trophic factor withdrawal and ischemic brain injury. *J. Exp. Med.* **185**, 933–940 (1997).
- Fernandes-Alnemri, T. et al. *In vitro* activation of CPP32 and Mch3 by Mch4, a novel human apoptotic cysteine protease containing two FADD-like domains. *Proc. Natl Acad. Sci. USA* **93**, 7464–7469 (1996).
- Nunez, G., Benedict, M. A., Hu, Y. & Inohara, N. Caspases: the proteases of the apoptotic pathway. *Oncogene* **17**, 3237–3245 (1998).
- Slee, E. A., Adrain, C. & Martin, S. J. Serial killers: ordering caspase activation events in apoptosis. *Cell Death Differ.* **6**, 1067–1074 (1999).
- Aliprantis, A. O., Yang, R. B., Weiss, D. S., Godowski, P. & Zychlinsky, A. The apoptotic signaling pathway activated by Toll-like receptor-2. *EMBO J.* **19**, 3325–3336 (2000).
- Jung, D. Y. et al. TLR4, but not TLR2, signals autoregulatory apoptosis of cultured microglia: a critical role of IFN-beta as a decision maker. *J. Immunol.* **174**, 6467–6476 (2005).
- Kuno, R. et al. Autocrine activation of microglia by tumor necrosis factor-alpha. *J. Neuroimmunol.* **162**, 89–96 (2005).
- Storz, P., Doppler, H. & Toker, A. Protein kinase Cdelta selectively regulates protein kinase D-dependent activation of NF- κ B in oxidative stress signaling. *Mol. Cell. Biol.* **24**, 2614–2626 (2004).
- Vancurova, I., Miskolci, V. & Davidson, D. NF- κ B activation in tumor necrosis factor α -stimulated neutrophils is mediated by protein kinase C δ . Correlation to nuclear I κ B α . *J. Biol. Chem.* **276**, 19746–19752 (2001).
- Reyland, M. E., Anderson, S. M., Matassa, A. A., Barzen, K. A. & Quissell, D. O. Protein kinase C δ is essential for etoposide-induced apoptosis in salivary gland acinar cells. *J. Biol. Chem.* **274**, 19115–19123 (1999).
- Czlonkowska, A., Kohutnicka, M., Kurkowska-Jastrzebska, I. & Czlonkowski, A. Microglial reaction in MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) induced Parkinson's disease mice model. *Neurodegeneration* **5**, 137–143 (1996).
- Aarli, J. A. Role of cytokines in neurological disorders. *Curr. Med. Chem.* **10**, 1931–1937 (2003).
- Gonzalez-Scarano, F. & Baltuch, G. Microglia as mediators of inflammatory and degenerative diseases. *Annu. Rev. Neurosci.* **22**, 219–240 (1999).
- Jordan, J., Segura, T., Brea, D., Galindo, M. F. & Castillo, J. Inflammation as therapeutic objective in stroke. *Curr. Pharm. Des.* **14**, 3549–3564 (2008).
- Lenzlinger, P. M., Morganti-Kossmann, M. C., Laurer, H. L. & McIntosh, T. K. The duality of the inflammatory response to traumatic brain injury. *Mol. Neurobiol.* **24**, 169–181 (2001).
- Allan, S. M. & Rothwell, N. J. Inflammation in central nervous system injury. *Phil. Trans. R. Soc. Lond. B* **358**, 1669–1677 (2003).
- Karatas, H. et al. A nanomedicine transports a peptide caspase-3 inhibitor across the blood-brain barrier and provides neuroprotection. *J. Neurosci.* **29**, 13761–13769 (2009).
- Bilsland, J. & Harper, S. Caspases and neuroprotection. *Curr. Opin. Investig. Drugs* **3**, 1745–1752 (2002).
- Friedlander, R. M. Apoptosis and caspases in neurodegenerative diseases. *N. Engl. J. Med.* **348**, 1365–1375 (2003).
- Le, D. A. et al. Caspase activation and neuroprotection in caspase-3-deficient mice after *in vivo* cerebral ischemia and *in vitro* oxygen glucose deprivation. *Proc. Natl Acad. Sci. USA* **99**, 15188–15193 (2002).
- Joseph, B. et al. p57(Kip2) cooperates with Nurr1 in developing dopamine cells. *Proc. Natl Acad. Sci. USA* **100**, 15619–15624 (2003).
- Bocchini, V. et al. An immortalized cell line expresses properties of activated microglial cells. *J. Neurosci. Res.* **31**, 616–621 (1992).
- Giulian, D. & Baker, T. J. Characterization of ameboid microglia isolated from developing mammalian brain. *J. Neurosci.* **6**, 2163–2178 (1986).
- Li, J. Y. et al. Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nature Med.* **14**, 501–503 (2008).
- Joseph, B. et al. Mitochondrial dysfunction is an essential step for killing of non-small cell lung carcinomas resistant to conventional treatment. *Oncogene* **21**, 65–77 (2002).
- Rite, I., Machado, A., Cano, J. & Venero, J. L. Blood-brain barrier disruption induces *in vivo* degeneration of nigral dopaminergic neurons. *J. Neurochem.* **101**, 1567–1582 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Gorman, O. Hermanson, M. Malewicz, S. Orrenius, T. Panaretakis and B. Zhivotovsky for discussion, and L. Hjortsberg, M. Reyland and S. Ceccatelli for providing us with reagents. M. Carballo, J.L. Ribas, A. Fernández and B. Haraldsson provided qualified technical support. This work has been supported by grants from the Spanish Ministerio de Ciencia y Tecnología (SAF2006-04119 and 2009-13778), the Swedish Research Council, the Parkinson Foundation of Sweden, the Swedish Alzheimer Foundation and the Swedish Cancer Society. M.A.B., T.D. and P.B. are members of Neurofortis and Bagadilico, both of which are research environments sponsored by the Swedish Research Council.

Author Contributions M.A.B. performed all the experiments except as otherwise noted. qPCR was performed by A.G.-Q. and E.K. J.L.V. and J.C. collaborated in doing surgery and further dissecting the animal brains. M.A.B. and T.D. performed primary cell culture experiments and cytokine analysis. E.K. collaborated in performing the caspase activity assay. B.J. and E.K. collaborated in performing FACS. B.J. collaborated also in the confocal imaging analysis. E.E. did the neuropathology of the individuals with PD and AD and the controls. A.P. prepared tissue and participated in the morphological assessment of human brain specimens. N.H. and P.B. were involved in study design. M.A.B., J.L.V. and B.J. designed the study, analysed and interpreted the data. All authors discussed the results and commented on or edited the manuscript. The first draft of the paper was written by B.J. J.L.V. and B.J. share senior authorship of the paper. T.D. and E.K. share second authorship.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.J. (bertrand.joseph@ki.se) or J.L.V. (jlvenero@us.es).

METHODS

Reagents. LPS (from *Escherichia coli*, serotype 026:B6; Sigma), staurosporine (STS; Sigma), MPTP (Sigma), agonist anti-Fas monoclonal antibody clone CH11 (MBL), agonist anti-TNF- α (R&D Systems), the caspase-3/7 inhibitor benzyloxycarbonyl-Asp(OMe)-Glu(OMe)-Val-Asp(OMe)-fluoromethylketone (DEVD-fmk), the caspase-8 inhibitor benzyloxycarbonyl-Ile-Glu(OMe)-Thr-Asp(OMe)-fluoromethylketone (IETD-fmk), the caspase-6 inhibitor benzyloxycarbonyl-Val-Glu(OMe)-Ile-Asp(OMe)-fluoromethylketone (VEID-fmk), the caspase-1 inhibitor benzyloxycarbonyl-Tyr-Val-Ala-Asp(OMe)-fluoromethylketone (YVAD-fmk; MP Biomedicals) and the PKC- δ inhibitor Rottlerin (Calbiochem) were used in this study. Plasmids encoding PKC- δ and EGFP-PKC- δ D327A were gifts of L. Hjörstberg and M. Reyland respectively. ON-TARGET plus SMARTpools siRNAs were purchased from Dharmacon (Supplementary Table 3).

Animals and surgery. Animals used in this study were obtained from the Center of Production and Animal Experimentation (Estartinas, Spain) and NMRI (Charles River, Germany). Experiments were performed in accordance with the Guidelines of the European Union Council (86/609/EU), following Spanish and Swedish regulations for the use of laboratory animals and approved by the Scientific Committee of the University of Seville and Lund University. Intraneural injections were made 5.8 mm anterior, 2.0 mm lateral and 8.0 mm ventral to the bregma in rat, and 3.1 mm anterior, \pm 1.2 mm lateral and 5.1 mm ventral from bregma in mouse.

Co-culture and neuronal cell death assay. Microglial BV2 and dopaminergic MN9D murine cells were cultured as described^{42,43}. MN9D dopaminergic neuronal cells were stained with CellTracker Green CMFDA (Invitrogen) before BV2 microglia cells transfected with caspase-3 siRNA, caspase-7 siRNA or non-targeting siRNA were plated on them. After 24 h, cells were treated with LPS and incubated for an additional 24 h. They were then fixed with 4% paraformaldehyde and stained with 0.1 mg ml⁻¹ Hoechst for quantification of neuronal cell death.

Primary cultures. Primary mouse microglial cells were plated for at least 48 h before the experiments (ethical permit M302-09). Cells were pretreated with inhibitors (20 μ M DEVD-fmk, 20 μ M IETD-fmk and 2 μ M Rottlerin 1 h before with 100 ng ml⁻¹ LPS treatment)⁴⁴.

Human brain. Human brain tissues from patients with PD of 5, 9 and 15 years' duration, respectively, and age-matched controls were used in this study⁴⁵. The region investigated was the anterior mesencephalon covering the substantia nigra. Frontal cortex from patients with AD of 4, 10 and 14 years' duration and controls were also used (Regional Ethical Review Board, Lund (Sweden) 2009-646/2010-25). The patients with dissimilar disease duration exhibited different degrees of severity of brain disease, reflecting different stages of the degenerative process. All sections were stained with haematoxylin-erythrosin and with antibodies against caspase 3, 8 and microglia, CD68 (see below). They were microscopically reviewed for verification of pathology. Before the investigation, the entire collection of brain sections, 15–20 per individual including the mesencephalic section, were subjected to a neuropathological whole-brain analysis for clinical diagnostic purposes, according to routine procedures at the Department of Pathology, Division of Neuropathology, Lund University Hospital. The project procedures involving human brain tissue were approved by the Regional Ethical Review Board in Lund, Sweden.

Immunofluorescence and laser scanning confocal microscopy. Paraformaldehyde-fixed cells were blocked in PBS/3% goat serum/0.3% triton X-100 and incubated

with the indicated primary (4 °C, overnight) and AlexaFluor 488 or 594 conjugated anti-IgG used as secondary antibodies (room temperature, ~22 °C, 1 h; Molecular Probes). Nuclei were counterstained with DRAQ5TM (1 μ M, Alexis), Hoechst (1 μ g ml⁻¹, Molecular Probes) or DAPI (1 μ g ml⁻¹, Molecular Probes). Alexa-555-conjugated cholera toxin B (CTB; 10 μ g ml⁻¹; Molecular Probes) was used to stain lipid rafts on the plasma membrane. Protein subcellular localization was analysed under a Zeiss 510 META or Leica TCS-SP2 confocal laser scanning microscope. The nuclear translocation of the NF- κ B p65 subunit was quantified using multicolour three-dimensional plug-in from Leica Confocal software and measured as the percentage of co-localization of DRAQ5TM with NF- κ B p65 inside the nucleus.

FACS analysis. Quantification of cells with cleaved caspase-3 and cleaved caspase-8 was performed with a FACSCalibur flow cytometer (Becton Dickinson) using standard procedures. Analysis of data was performed using Cell Quest software⁴⁶.

Immunohistochemistry and immunohistochemistry. Rats and mice were perfused through the heart under deep anaesthesia with 4% paraformaldehyde/PBS, pH 7.4. Brains were removed, cryoprotected in sucrose and frozen in isopentane at -15 °C; serial coronal sections (25- and 30- μ m sections, for rat and mice, respectively) covering the striatum and the substantia nigra were cut with a cryostat and mounted on gelatin-coated slides⁴⁷. Sections were incubated with the indicated primary antibodies. After three washes, sections were incubated with biotinylated horse anti-mouse or goat anti-rabbit IgG (Vector) followed by an incubation with ExtrAvidin-Peroxidase solution (Sigma) and for immunofluorescence by a fluorescein isothiocyanate (FITC)-conjugated anti-rabbit and Texas Red anti-mouse antibody (Vector). The peroxidase was visualized with a standard diaminobenzidine/hydrogen reaction for 5 min. For paraffin-embedded human tissue material, sections (5 μ m) were mounted on capillary glass slides (DAKO). Sections were microwaved pre-treated in 10 mM citrate buffer pH 6.0 for 15 min at 800 W for antigen retrieval. An automated immunostainer (TechMateTM 500 Plus, DAKO) was used for the staining procedure using DAKO ChemMate Kit Peroxidase/3-3'-diaminobenzidine. The indicated primary antibodies were used.

Quantification of microglial population in animal models. Reactive and resident microglial cells were counted in LPS-injected rats and MPTP-exposed mice detected by Iba1 immunohistochemistry based on morphological features. For each animal, eight sections covering the entire antero-posterior ventral mesencephalon were analysed. For each section, four photographs were taken at \times 20 magnification (two for each substantia nigra) and microglial cells were counted with the analySIS software.

Quantification of the striato-nigral dopaminergic system. Tyrosine hydroxylase immunohistochemistry in the striatum of mice intoxicated with MPTP was quantified using a computer-assisted software (analySIS). For the quantification, five striatal sections from each condition, processed under identical experimental conditions, were scanned at high resolution. The striatal region was delineated and its optical density measured based upon a calibrated grey scale. Quantification of tyrosine-hydroxylase-positive cells in the substantia nigra was performed according to a modified stereological approach using the Olympus CAST-Grid system. The area of the substantia nigra region was estimated using the principle of Cavalieri. All data were collected while blind to experimental treatment and expressed as number of neurons per substantia nigra.

Crystal structure of metarhodopsin II

Hui-Woog Choe^{1,2*}, Yong Ju Kim^{1*}, Jung Hee Park^{1*}, Takefumi Morizumi^{1†}, Emil F. Pai³, Norbert Krauß⁴, Klaus Peter Hofmann^{1,5}, Patrick Scheerer¹ & Oliver P. Ernst^{1†}

G-protein-coupled receptors (GPCRs) are seven transmembrane helix (TM) proteins that transduce signals into living cells by binding extracellular ligands and coupling to intracellular heterotrimeric G proteins (Gαβγ)¹. The photoreceptor rhodopsin couples to transducin and bears its ligand 11-*cis*-retinal covalently bound via a protonated Schiff base to the opsin apoprotein². Absorption of a photon causes retinal *cis/trans* isomerization and generates the agonist all-*trans*-retinal *in situ*. After early photoproducts, the active G-protein-binding intermediate metarhodopsin II (Meta II) is formed, in which the retinal Schiff base is still intact but deprotonated. Dissociation of the proton from the Schiff base breaks a major constraint in the protein and enables further activating steps, including an outward tilt of TM6 and formation of a large cytoplasmic crevice for uptake of the interacting C terminus of the Gα subunit^{3–5}. Owing to Schiff base hydrolysis, Meta II is short-lived and notoriously difficult to crystallize. We therefore soaked opsin crystals with all-*trans*-retinal to form Meta II, presuming that the crystal's high concentration of opsin in an active conformation (Ops*)^{6,7} may facilitate all-*trans*-retinal uptake and Schiff base formation. Here we present the 3.0 Å and 2.85 Å crystal structures, respectively, of Meta II alone or in complex with an 11-amino-acid C-terminal fragment derived from Gα (GαCT2). GαCT2 binds in a large crevice at the cytoplasmic side, akin to the binding of a similar Gα-derived peptide to Ops* (ref. 7). In the Meta II structures, the electron density from the retinal ligand seamlessly continues into the Lys 296 side chain, reflecting proper formation of the Schiff base linkage. The retinal is in a relaxed conformation and almost undistorted compared with pure crystalline all-*trans*-retinal. By comparison with early photoproducts we propose how retinal translocation and rotation induce the gross conformational changes characteristic for Meta II. The structures can now serve as models for the large GPCR family.

To obtain crystals of Meta II alone or in complex with the GαCT2 peptide, we followed previous approaches using low pH to obtain crystals of opsin in the active Ops* conformation^{6,7} for subsequent soaking with all-*trans*-retinal. GαCT2 was derived from the transducin Gα C terminus and contained two substitutions which improved crystal diffraction quality compared with crystals containing the native peptide fragment (see Methods). Reversible formation of Meta II is indicated by (1) the 380-nm absorption obtained before and after dissolving the all-*trans*-retinal-soaked opsin or opsin–GαCT2 crystals, respectively, (2) formation of an opsin-retinal Schiff base bond as determined by acid denaturation of the reconstituted pigments (Supplementary Fig. 1), (3) the X-ray crystallographic evidence for the Schiff base bond between all-*trans*-retinal and Lys 296^{7,43} of opsin (see below), and (4) the high activity of the reconstituted pigment towards the G protein (Supplementary Fig. 2). In addition, soaking of the opsin crystals with all-*trans*-retinal for an optimized soaking time had only little effect on the diffraction quality of the opsin crystals,

indicating that the Ops* conformation present in the crystal is compatible with the all-*trans*-retinal agonist. In contrast, soaking of opsin crystals with the inverse agonist 11-*cis*-retinal led to a reddish colour indicative of rhodopsin formation. Soaking was accompanied by severe loss of crystal diffraction quality, indicating disintegration of crystal packing, consistent with the expected larger conformational changes induced in the protein by 11-*cis*-retinal.

The structures of Meta II and Meta II–GαCT2 were solved at 3.0 and 2.85 Å resolution, respectively, by molecular replacement techniques using opsin in the Ops* conformation (PDB accession 3DQB) as search model. For crystallization, data collection, structure determination and refinement statistics, see Methods and Supplementary Table 1. The structures comprise amino acids 1–326 of the opsin protein moiety, lacking 22 C-terminal amino acids, which are not resolved in the corresponding electron density maps, presumably because of high mobility^{6,7}. Each model also includes one molecule all-*trans*-retinal, oligosaccharyl chains at Asn 2^{NT} (for superscripts see Methods Summary) and Asn 15^{NT}, several water and two *n*-octyl-β-D-glucopyranoside molecules, and in addition in the case of Meta II–GαCT2, the 11-amino-acid GαCT2 peptide. The models further comprise one molecule palmitic acid covalently bound to Cys 323^{H8}; a second palmitic acid residue at Cys 322^{H8} is lacking in each model because of weak electron density. The overall structure of Meta II and Meta II–GαCT2, comprising seven TMs connected by cytoplasmic and extracellular loops (CL1–CL3 and EL1–EL3) which are followed by the cytoplasmic helix H8, is similar to the previously solved Ops* and Ops*–GαCT structures^{6,7}, but quite different from inactive rhodopsin^{8–10}. The Cα root mean squared deviation (r.m.s.d.) values in Å are 0.51 (Ops*/Meta II), 0.44 (Ops*–GαCT/Meta II–GαCT2), 0.35 (Meta II/Meta II–GαCT2) and 2.85 (Rhodopsin/Meta II) (Fig. 1 and Supplementary Fig. 3)^{6,7,9}. Both Meta II and Meta II–GαCT2 display the open cytoplasmic receptor surface: due to a rotational tilt of TM6 and motion of TM5 a crevice in the cytoplasmic surface is created into which the Gα C terminus or its mimic, the GαCT2 peptide, can bind^{4,5,7}.

The Meta II structures now provide insight into retinal–protein interactions in the ligand binding site. The initial electron density map calculated with omitted retinal ligand could readily be interpreted with the relaxed all-*trans*-retinal conformation represented by the model obtained from the Cambridge Structural Database (reference code TRETAL02; Supplementary Figs 4 and 5). There is strong electron density connecting the retinal into the side chain of Lys 296^{7,43} consistent with the presence of a Schiff base bond between carbon atom C15 of retinal and the ε-amino group of Lys 296^{7,43} (Fig. 2 and Supplementary Fig. 4). The refinement of the Meta II structures only slightly changed the initial conformation of all-*trans*-retinal, indicating that retinal in a relaxed conformation is bound in the ligand binding site of Ops*, in agreement with spectroscopic studies on Meta II (ref. 11). Both structures, Meta II and Meta II–GαCT2, show an almost identical retinal conformation (Supplementary Fig. 5). The

¹Institut für Medizinische Physik und Biophysik - CC2, Charité - Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin, Germany. ²Department of Chemistry, College of Natural Science, Chonbuk National University, 561-756 Chonju, South Korea. ³Departments of Biochemistry, Molecular Genetics and Medical Biophysics, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada. ⁴Queen Mary University of London, School of Biological and Chemical Sciences, London E1 4NS, UK. ⁵Zentrum für Biophysik und Bioinformatik, Humboldt-Universität zu Berlin, Invalidenstrasse 42, D-10115 Berlin, Germany. [†]Present addresses: Department of Biochemistry, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada (T.M.); Departments of Biochemistry and Molecular Genetics, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada (O.P.E.).

*These authors contributed equally to this work.

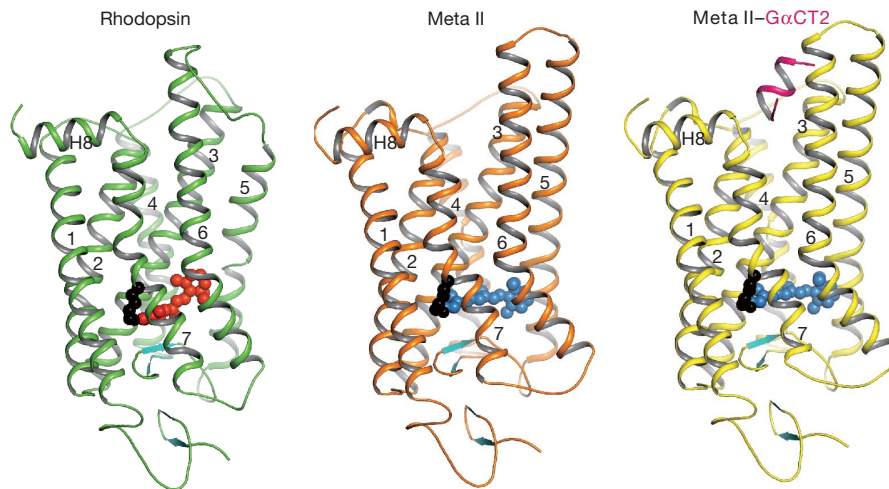


Figure 1 | Structures of inactive rhodopsin, active Meta II and Meta II in complex with a Ga fragment. Side view of cartoon model of rhodopsin (PDB accession 1U19), Meta II and Meta II-GαCT2, that is, Meta II in complex with an 11-amino-acid C-terminal peptide derived from the transducin Gα subunit, representing the key binding site on the heterotrimeric G protein. Inactive rhodopsin consists of the apoprotein opsin in its inactive conformation and the inverse agonist 11-*cis*-retinal (shown as red spheres), which is covalently bound to Lys 296 (shown as black spheres) via a Schiff base. In Meta II, the agonist all-

trans-retinal (shown as blue spheres) is covalently linked to Lys 296. The G-protein-interacting conformation of Meta II (right panel) is identified by the bound Gα fragment (shown in purple). Transmembrane helices (numbered 1–7) are followed by a cytoplasmic helix H8. The Meta II models lack 22 C-terminal residues, the C terminus of rhodopsin is not shown. The extracellular domain contains four β-strands (coloured in cyan). Glycosylation at Asn 2 and Asn 15 and palmitoylation at Cys 322 and Cys 323 (at the end of H8) are not shown.

binding site for all-*trans*-retinal appears to be preformed in Ops* because the presence of retinal causes only a small adjustment of side chains (Trp 265^{6,48}, Tyr 268^{6,51}, Met 207^{5,42}, His 211^{5,46} and Glu 122^{3,37}) while the Lys 296^{7,43} side chain, which is more flexible in ligand-free Ops* (ref. 6), becomes ordered due to its linkage with retinal (Supplementary Fig. 6). The Meta II-GαCT2 structure reveals a potential

hydrogen bonding network in which a water molecule links the Schiff base nitrogen with the side chains of Ser 186^{EL2} and Glu 181^{EL2}, the latter hydrogen-bonded to the side chain of Tyr 268^{6,51} (Fig. 2 and Supplementary Fig. 7). A further water molecule constrains the main chain carbonyl of Glu 181^{EL2}, the side chain of Ser 186^{EL2} and the NH-groups of Ser 186^{EL2} and Cys 185^{EL2}. From Cys 185^{EL2} the hydrogen bonding

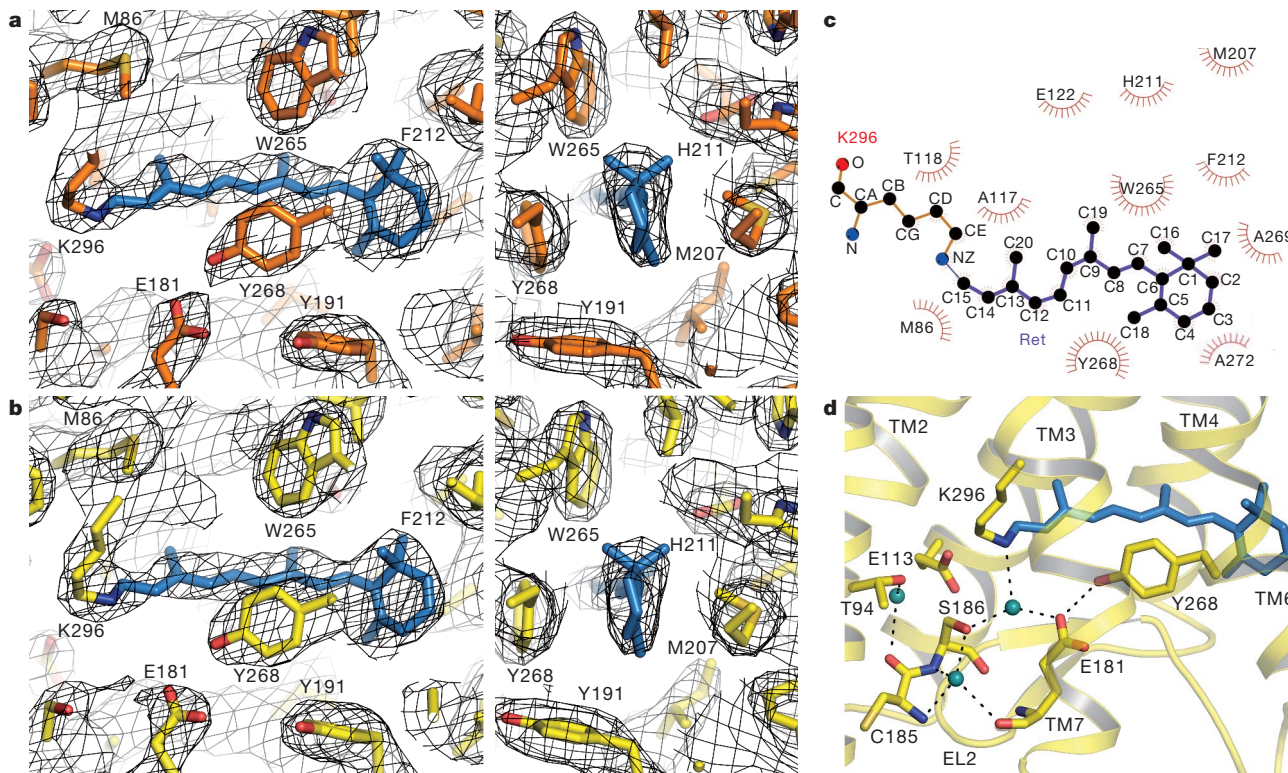


Figure 2 | Retinal binding pocket of Meta II. **a, b**, Retinal and neighbouring residues shown as stick model for Meta II (orange) and Meta II-GαCT2 (yellow), respectively, with $2F_o - F_c$ electron density map contoured at 1.0σ (grey mesh). View from within the membrane, right panel rotated 90° with respect to the left. Retinal is shown as blue stick model. **c**, Hydrophobic

interactions between all-*trans*-retinal and residues in the ligand binding pocket of the opsin protein moiety. Residues with closest distances less than 4 Å are considered to be in van der Waals contact. **d**, Potential hydrogen bonding interactions around retinal in Meta II-GαCT2. Water molecules are shown as blue spheres.

network extends via a water molecule to the side chain of Thr 94^{2,61}. In the inactive rhodopsin conformation, a different network is found in which the Schiff base nitrogen is within hydrogen-bonding distance to the side chain of Glu 113^{3,28} from which the network can extend to Thr 94^{2,61} and via Ser 186^{EL2} to Glu 181^{EL2} (refs 9, 10). Together Glu 113^{3,28} and Glu 181^{EL2} form the counterion for the protonated Schiff base in inactive rhodopsin. From Fourier transform infrared (FTIR) studies on rhodopsin activation it was predicted that the main contribution to this complex counterion changes from Glu 113^{3,28} to Glu 181^{EL2} during the transition to Meta I (ref. 12). This is also reflected in the Meta II structure. Relative to rhodopsin, the distance between the Schiff base and the side-chain carboxyl group of Glu 113^{3,28} increases from 3.5 to 5.3 Å, whereas for Glu 181^{EL2} the corresponding distance decreases from 7.1 to 5.0 Å.

The Meta II-G α CT2 structure reveals additional structural water molecules next to amino acids in conserved motifs, extending the hydrogen-bonding network from Asn 302^{7,49} of the NPxxY(x)_{5,6}F motif to Glu 134^{3,49} of the E(D)RY motif. The lack of an interaction between Glu 134^{3,49} and Arg 135^{3,50} in Meta II-G α CT2 enables Arg 135^{3,50} to bind G α CT2 by forming a hydrogen bond to the backbone carbonyl oxygen of Val 347 of G α CT2. TM2 and TM7 are linked via a water molecule between Ser 298^{7,45} and highly conserved residues Asp 83^{2,50} and Asn 302^{7,49}. Structural water molecules are conserved in GPCRs¹³ and changes in radiolytic hydroxyl radical labelling have indicated activation-induced alterations in the local structures in which they are embedded¹⁴. Our data provide further insight into how water molecules are rearranged in the course of receptor activation (Supplementary Fig. 7).

In contrast to rhodopsin^{8–10}, but like Ops* (refs 6, 15), the Meta II structures show two openings of the retinal binding pocket (Supplementary Fig. 8). The opening between TM1 and TM7 is caused by a rotamer change of Phe 293^{7,40} relative to rhodopsin, which occurs in the course of a rotational rearrangement of the extracellular part of TM7 below the retinal attachment site Lys 296^{7,43} (Supplementary Fig. 9). The opening between TM5 and TM6 arises mainly from side-chain changes of Phe 208^{5,43}, Phe 212^{5,47} and Phe 273^{6,56} which are also observed in the Ops* structure⁶. The location of retinal in its binding site fulfils most distance constraints obtained from dipolar-assisted rotational resonance (DARR) NMR spectroscopy (Supplementary Table 2)^{16,17}. A few larger distance deviations indicate that perhaps different Meta II sub-states were observed in the NMR and X-ray experiments (see Scheme 1 in the Supporting Information).

Meta II shows, like Ops*, two gross conformational changes with respect to rhodopsin (Supplementary Fig. 3)^{6,7}. Motion of TM5 towards TM6 is linked to a rearrangement of the TM3-TM5 hydrogen bonding network between Glu 122^{3,37}, Trp 126^{3,41} and His 211^{5,46}. In Meta II, the side chains of Glu 122^{3,37} and His 211^{5,46} are close enough to form a hydrogen bond (Supplementary Figs 6 and 7). The second extensive conformational change involves TM6, which undergoes a rotational tilt resulting in a motion of its cytoplasmic end away from TM3 (Figs 3 and 4). For the corresponding structural rearrangement of residues in the conserved motifs E(D)RY (in TM3) and NPxxY(x)_{5,6}F (in TM7-H8) see the Supplementary Discussion.

Unexpectedly, all-*trans*-retinal in the reversibly formed Meta II appears to be rotated around its long axis compared with the dark state^{8–10}, and the early photoproducts Batho¹⁸ and Lumi¹⁹ (Fig. 3). The significant electron density observed for the methyl groups of the retinal polyene chain and β -ionone ring gives a preference for modelling retinal rotated approximately upside-down relative to Lumi (Fig. 3a) and non-satisfactory results when only the polyene chain was rotated but the β -ionone ring was kept as in Lumi (6-*s-trans* conformation). Retinal rotation is unlikely to occur with the packing seen in the Meta II crystal (Fig. 2c), but small TM movements in the receptor activation process, which can occur as early as in Meta I²⁰, may well provide the necessary space. The Meta II structure does neither reflect the proposed ‘toggle switch’ rotational change of Trp 265^{6,48}

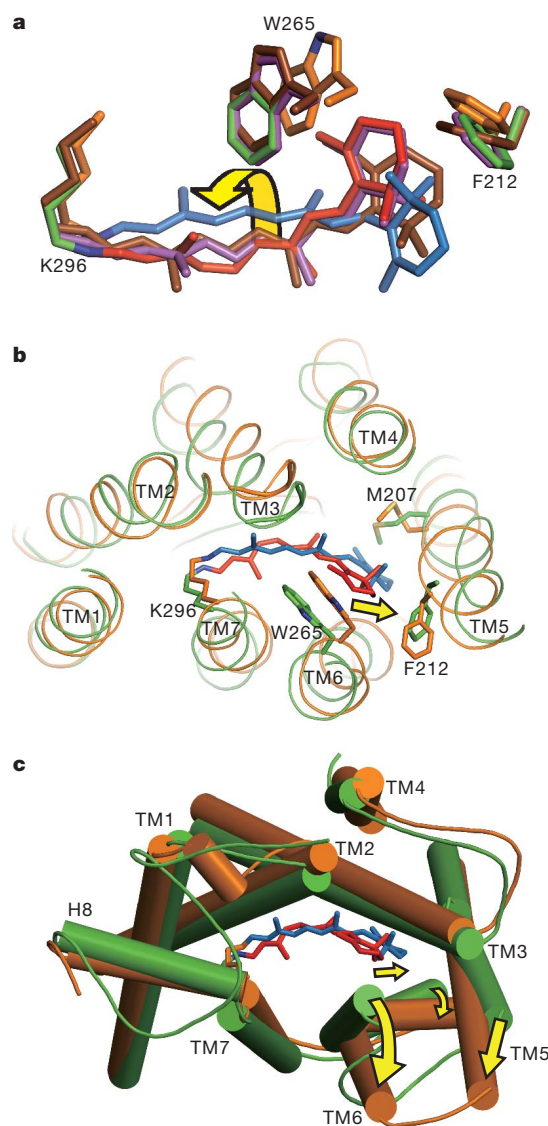


Figure 3 | Superposition of rhodopsin, Batho, Lumi and Meta II. **a**, View from within the membrane on superposition of retinal in rhodopsin (PDB 1U19), Batho (PDB 2G87), Lumi (PDB 2HPY) and Meta II. Amino acid side chains and retinal are shown as stick models. Protein/retinal are coloured green/red (rhodopsin), orange/blue (Meta II), purple/purple (Batho) and brown/brown (Lumi), respectively. The arrow indicates rotation of retinal from Lumi to Meta II. **b**, **c**, Superposition of rhodopsin and Meta II viewed from the cytoplasmic surface. Helices are shown as cartoon or cylinder models, retinal and amino acid side chains as stick models. For clarity, cytoplasmic ends of helices are clipped in **b**. Colours as in **a**. Elongation of retinal and interaction of the β -ionone-ring moiety with TM5/TM6 induces a rigid body-type rotational tilt of TM6 and motion of TM5 (indicated by arrows).

side chain^{17,21–23} nor the displacement of loop EL2 from the retinal binding site¹⁶. An explanation would be that these structural changes are transient and involved in switching to rather than maintaining the active conformation (see Supplementary Discussion and Supplementary Fig. 3). The Meta II structure also shows a positively twisted 6-*s-cis* conformation regarding the bond between carbon atoms C6 and C7 where retinal's β -ionone ring and polyene chain are connected. A negatively twisted 6-*s-cis* conformation as in the rhodopsin dark state, Batho and Lumi, was predicted for Meta I and Meta II by NMR spectroscopy^{17,24–26}, although flexibility of retinal even in the inactive receptor

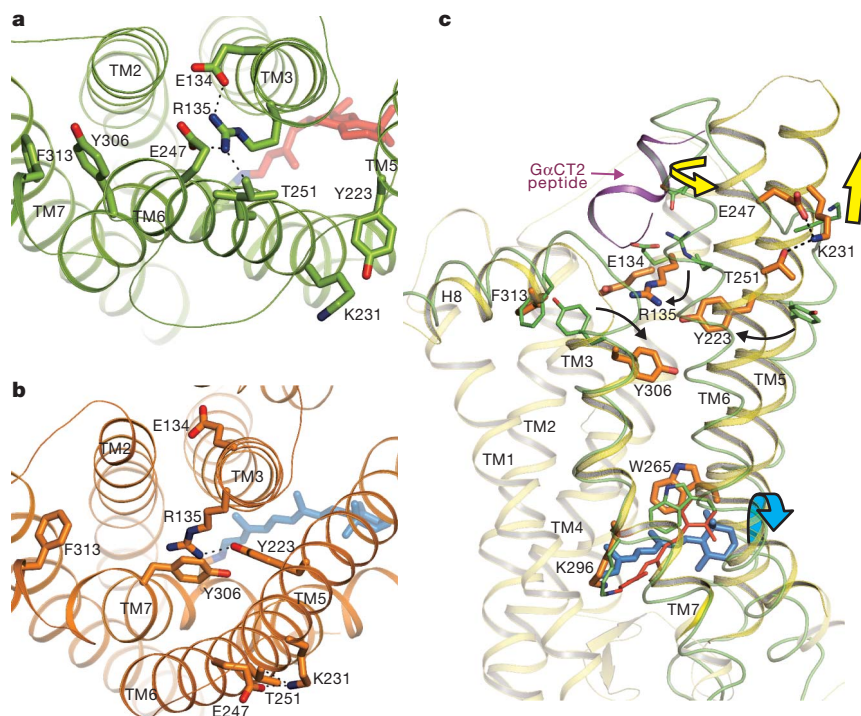


Figure 4 | Conserved E(D)RY and NPxxY(x)_{5,6}F regions. **a**, Rhodopsin and **b**, Meta II viewed from the cytoplasmic surface. Residues in the conserved E(D)RY and NPxxY(x)_{5,6}F motifs in TM3 and TM7/H8, respectively, constrain rhodopsin (green) in the inactive state. A hydrogen bonded network termed ‘ionic lock’ is formed between TM3 and TM6 including Glu 134^{3,49}, Arg 135^{3,50}, Glu 247^{6,30} and Thr 251^{6,34}. Tyr 306^{7,53} and Phe 313^{7,60} show an electrostatic interaction and are part of a functional microdomain between TM7 and H8. In Meta II and Meta II–GαCT2, the conserved residues form a different set of interactions. Glu 247^{6,30}–Lys 231^{5,66}–Thr 251^{6,34} and Tyr 223^{5,58}–Arg 135^{3,50}

state may allow negatively and positively twisted 6-*s-cis* conformations²⁷. An artificial pigment, in which ring rotation about the C6–C7 bond is inhibited, was reported to show lower but still substantial activity towards the G protein²⁸. The positive C5–C6–C7–C8 dihedral angle may thus be preferred but not mandatory to form an active state. Generally, a special status of the retinal in the active Meta II form fits well to previous FTIR spectroscopic results, which have shown that photo-induced back reactions are different between active Meta II and its inactive precursors (Supplementary Discussion).

The X-ray analysis of the reversibly formed Meta II has shown that the ligand fits readily into the binding pocket of Ops* and displays the same relaxed conformation as in pure retinal crystals. Ops* remains in its overall structure almost unchanged when the all-*trans*-retinal ligand binds, and the cytoplasmic G protein interaction domain of Ops* and Meta II carry the same insignia of an active state. On the other hand, comparison of rhodopsin and Meta II structures indicates that the retinal must go through complex elongation and torsional motions of its polyene chain and of the β-ionone ring. The motions of the ligand are likely to correspond to identified intermediates in the light induced pathway. Further structural and spectroscopic analyses may teach us how retinal analogues are less efficiently interrogated by the protein (ref. 29 and references therein), and may open insights into partial agonism and β-arrestin biased ligands at GPCRs. A more general lesson to be learned from this work on rhodopsin is that for reaching the active GPCR state, the receptor can make use of the conformational flexibility of the ligand and the variability of its interaction with the binding site.

METHODS SUMMARY

Opsin was extracted from bovine rod outer segment disc membranes using *n*-octyl-β-D-glucopyranoside and *n*-dodecyl-β-D-maltopyranoside. Solubilized

opsin was crystallized alone or in the presence of GαCT2 peptide (NH₂-ILENLKDVGLF-COOH; Gα_{340–350}(K341L,C347V)) by hanging-drop vapour diffusion in a mixture of ammonium sulphate and sodium acetate buffer at pH 5.0–5.8. Hundreds of opsin and opsin–GαCT2 crystals grown within 5 days were systematically soaked with all-*trans*-retinal for up to 30 h in the dark to reconstitute Meta II and Meta II–GαCT2, respectively. Crystals were cryoprotected in 10% trehalose and frozen in liquid nitrogen for screening and X-ray analysis at the synchrotron BESSY II, Berlin. The Meta II structures were solved by molecular replacement using opsin (PDB accession 3DQB) as search model.

The superscript denotes N terminus (NT; for example, Asn 2^{NT}), Helix 8 (H8; for example, Phe 313^{H8}), extracellular loop (EL; for example, Glu 181^{E12}), and GPCR numbering according to Ballesteros and Weinstein³⁰. The most conserved residue in a helix among rhodopsin-like GPCRs is designated x.50, where x is the helix number. All other residues on the same helix are numbered relative to this reference residue. Based on extensive sequence alignment, Smith grouped residues into signature-conserved, group-conserved or subfamily-specific¹⁷. Underlined superscript in GPCR numbering indicates GPCR signature residues (most conserved residues; Asn 55^{1.50}, Asp 83^{2.50}, Glu 134^{3.49}, Arg 135^{3.50}, Trp 161^{4.50}, Pro 215^{5.50}, Tyr 223^{5.58}, Trp 265^{6.48}, Pro 267^{6.50}, Asn 302^{7.49}, Pro 303^{7.50}, Tyr 306^{7.53}). Note that in ref. 17, based on a larger GPCR database, Leu 79 and Tyr 223 were attributed the most conserved residues on TM2 and TM5, respectively.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 September 2010; accepted 5 January 2011.

Published online 9 March 2011.

1. Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363 (2009).
2. Palczewski, K. G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.* **75**, 743–767 (2006).
3. Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl Acad. Sci. USA* **105**, 7439–7444 (2008).

4. Hofmann, K. P. *et al.* A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem. Sci.* **34**, 540–552 (2009).
5. Choe, H.-W., Park, J. H., Kim, Y. J. & Ernst, O. P. Transmembrane signaling by GPCRs: insight from rhodopsin and opsin structures. *Neuropharmacology* **60**, 52–57 (2011).
6. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H.-W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183–187 (2008).
7. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).
8. Palczewski, K. *et al.* Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745 (2000).
9. Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **342**, 571–583 (2004).
10. Li, J., Edwards, P. C., Burghammer, M., Villa, C. & Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343**, 1409–1438 (2004).
11. Siebert, F. Application of FTIR spectroscopy to the investigation of dark structures and photoreactions of visual pigments. *Isr. J. Chem.* **35**, 309–323 (1995).
12. Lücke, S. *et al.* The role of Glu181 in the photoactivation of rhodopsin. *J. Mol. Biol.* **353**, 345–356 (2005).
13. Angel, T. E., Chance, M. R. & Palczewski, K. Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proc. Natl Acad. Sci. USA* **106**, 8555–8560 (2009).
14. Angel, T. E., Gupta, S., Jastrzebska, B., Palczewski, K. & Chance, M. R. Structural waters define a functional channel mediating activation of the GPCR, rhodopsin. *Proc. Natl Acad. Sci. USA* **106**, 14367–14372 (2009).
15. Hildebrand, P. W. *et al.* A ligand channel through the G protein coupled receptor opsin. *PLoS ONE* **4**, e4382 (2009).
16. Ahuja, S. *et al.* Helix movement is coupled to displacement of the second extracellular loop in rhodopsin activation. *Nature Struct. Mol. Biol.* **16**, 168–175 (2009).
17. Smith, S. O. Structure and activation of the visual pigment rhodopsin. *Annu. Rev. Biophys.* **39**, 309–328 (2010).
18. Nakamichi, H. & Okada, T. Crystallographic analysis of primary visual photochemistry. *Angew. Chem. Int. Edn Engl.* **45**, 4270–4273 (2006).
19. Nakamichi, H. & Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103**, 12729–12734 (2006).
20. Ye, S. *et al.* Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. *Nature* **464**, 1386–1389 (2010).
21. Shi, L. *et al.* β 2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J. Biol. Chem.* **277**, 40989–40996 (2002).
22. Crocker, E. *et al.* Location of Trp265 in metarhodopsin II: implications for the activation mechanism of the visual receptor rhodopsin. *J. Mol. Biol.* **357**, 163–172 (2006).
23. Nygaard, R., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Schwartz, T. W. Ligand binding and micro-switches in 7TM receptor structures. *Trends Pharmacol. Sci.* **30**, 249–259 (2009).
24. Salgado, G. F. *et al.* Solid-state ^2H NMR structure of retinal in metarhodopsin I. *J. Am. Chem. Soc.* **128**, 11067–11071 (2006).
25. Ahuja, S. *et al.* 6-*s-cis* conformation and polar binding pocket of the retinal chromophore in the photoactivated state of rhodopsin. *J. Am. Chem. Soc.* **131**, 15160–15169 (2009).
26. Brown, M. F., Salgado, G. F. & Struts, A. V. Retinal dynamics during light activation of rhodopsin revealed by solid-state NMR spectroscopy. *Biochim. Biophys. Acta* **1798**, 177–193 (2010).
27. Lau, P. W., Grossfield, A., Feller, S. E., Pitman, M. C. & Brown, M. F. Dynamic structure of retinylidene ligand of rhodopsin probed by molecular simulations. *J. Mol. Biol.* **372**, 906–917 (2007).
28. Fujimoto, Y. *et al.* On the bioactive conformation of the rhodopsin chromophore: absolute sense of twist around the 6-*s-cis* bond. *Chem. Eur. J.* **7**, 4198–4204 (2001).
29. Knierim, B., Hofmann, K. P., Gartner, W., Hubbell, W. L. & Ernst, O. P. Rhodopsin and 9-demethyl-retinal analog: effect of a partial agonist on displacement of transmembrane helix 6 in class A G protein-coupled receptors. *J. Biol. Chem.* **283**, 4967–4974 (2008).
30. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G-protein coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Engelmann, C. Koch and B. Bauer for technical assistance, and F. Siebert and W. Hubbell for critically reading the manuscript. We are grateful to the European Synchrotron Radiation Facility (ESRF, Grenoble), D. von Stetten and A. Royant of the ID29S-Cryobench (ESRF, Grenoble) and U. Müller and the scientific staff of the BESSY-MX/Helmholtz Zentrum Berlin für Materialien und Energie at beamlines BL 14.1 and BL 14.2, where the data were collected, for continuous support. This work was supported by the DFG Sfb449 (to O.P.E.), Sfb740 (to O.P.E. and K.P.H.) and an Advanced Investigator ERC grant (to K.P.H.) and by the Canada Research Chairs Program (to E.F.P.). H.-W.C. gratefully acknowledges the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0002738) and CBNU funds for overseas research 2009. Y.J.K. thanks the Leibniz Graduate School of Molecular Biophysics, Berlin, for a scholarship.

Author Contributions H.-W.C., Y.J.K. and J.H.P. are joint first authors. H.-W.C., Y.J.K., J.H.P. performed preparation and crystallization of opsin/opsin–G α CT2. H.-W.C. performed the soaking experiment of both crystals. O.P.E. designed G α CT2. H.-W.C., Y.J.K., J.H.P., P.S., O.P.E. performed the data collection. Y.J.K., P.S., N.K. performed the structural analysis of Meta II, and J.H.P., P.S., E.F.P. performed the structural analysis of Meta II•G α CT2. T.M. performed the spectroscopic and biochemical analysis. H.-W.C., N.K., K.P.H., P.S., O.P.E. analysed data and H.-W.C., K.P.H., O.P.E. wrote the paper with contributions from all authors.

Author Information Atomic coordinates and structure factors for the reported structure have been deposited in the Protein Data Bank with the accession codes 3PQR and 3PXO. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.-W.C. (hwchoe@bnu.ac.kr), K.P.H. (klaus_peter.hofmann@charite.de) or O.P.E. (oliver.ernst@utoronto.ca).

METHODS

Crystallization. Opsin was prepared and crystallized by hanging drop vapour diffusion at 277 K essentially as described^{6,7}. The solubilization buffer contained a mixture of 1% *n*-octyl- β -D-glucopyranoside and 0.02% *n*-dodecyl- β -D-maltopyranoside or *n*-octyl- β -D-glucopyranoside alone. The precipitant solution was 3.0–3.4 M $(\text{NH}_4)_2\text{SO}_4$ in 0.1 M sodium acetate buffer (pH 5.0–5.8). A synthetic peptide was used for co-crystallization of opsin with a C-terminal fragment derived from the α -subunit of the G protein transducin. The peptide G α CT2 (sequence NH₂-ILENLKDVGLF-COOH) is a derivative of the high-affinity G α CT peptide³¹ and has a similar affinity. Crystals appeared within 2 days and grew further for 3 days to reach dimensions 0.2 mm \times 0.3 mm \times 0.3 mm. All-*trans*-retinal was purchased from Sigma and further purified by normal phase HPLC using a LiChrosorb Si 60 (5 μ m) 300 mm \times 20 mm column with 5% diethylether in heptane. Opsin crystals were soaked in the dark with all-*trans*-retinal for up to 30 h to reconstitute Meta II. Crystals were then transferred under orange light to cryoprotectant consisting of 10% trehalose in crystallization buffer and immediately frozen in liquid nitrogen.

Structure analysis. Diffraction data collection was performed at 100 K using synchrotron X-ray sources at BESSY II, Berlin, Germany, and ESRF, Grenoble, France. Best diffraction data were collected at beamline BL 14.2 at BESSY II, at $\lambda = 0.91842$ Å. The crystal to MX-225 CCD detector distance was fixed at 230 mm for Meta II and 260 mm for Meta II–G α CT2, respectively. The rotation increment for each frame was 0.5° with an exposure time of 6 s (Meta II) or 12 s (Meta II–G α CT2). All images were indexed, integrated and scaled using the XDS program package³² and CCP4 program SCALA³³. Meta II and Meta II–G α CT2 crystals belong to rhombohedral space group *H*32 ($a = 241.82$ Å, $b = 241.82$ Å, $c = 111.80$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ and $a = 241.55$ Å, $b = 241.55$ Å, $c = 109.87$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$, respectively). Supplementary Table 1 summarizes the statistics for crystallographic data collection and structural refinement.

Initial phases for Meta II and Meta II–G α CT2 were obtained by conventional molecular replacement protocol (rotation, translation, rigid body fitting) using opsin from the opsin–G α CT structure (PDB accession 3DQB) as initial search model. Molecular replacement was achieved using the CCP4 program PHASER³³ by first placing the opsin monomer (rotation function (RFZ): $Z = 14.5$; translation function (TFZ): $Z = 53.2$ for Meta II and rotation function (RFZ): $Z = 11.7$; translation function (TFZ): $Z = 63.5$ for Meta II–G α CT2, respectively; RFZ and TFZ as defined by PHASER). In subsequent steps, torsion angle molecular dynamics, simulated annealing using a slow-cooling protocol and a maximum likelihood target function, energy minimization, and *B*-factor refinement by the program CNS³⁴ were carried out in the resolution range 33.5–3.0 Å (Meta II) and 34.9–2.85 Å (Meta II–G α CT2). After the first round of refinement, the retinal chromophore was clearly visible in the electron density of both σ_A -weighted $2F_o - F_c$ maps, as well as in the σ_A -weighted simulated annealing omit density maps (Supplementary Fig. 4). Restrained, individual *B*-factors were refined and the crystal structure was finalized by the CCP4 program REFMAC5 and CCP4 (ref. 33). The final models have agreement factors R_{free} and R_{cryst} of 24.5% and 21.6% (Meta II), and 25.0% and 21.7% (Meta II–G α CT2), respectively. Manual rebuilding of the Meta II models and electron density interpretation were performed after each refinement cycle using the program COOT³⁵. Structure validation was performed with the programs PROCHECK³⁶ and WHAT_CHECK³⁷. Potential hydrogen bonds and van der Waals contacts were analysed using the programs HBPLUS³⁸ and LIGPLOT³⁹. All crystal structure superpositions of backbone alpha

carbon traces were performed using CCP4 program LSQKAB³³. All molecular graphics representations were created using PyMol⁴⁰.

Ultraviolet-visible spectroscopy. Ultraviolet-visible spectroscopy of Meta II was performed using a NanoDrop 1000 Spectrophotometer (Thermo Scientific) at room temperature and a CARY 50 Ultraviolet-visible Spectrophotometer (Varian) at 20 °C. Meta II crystals were washed thoroughly three times in crystallization buffer to reduce excess all-*trans*-retinal and were measured immediately in the same buffer. In some cases, Meta II crystals were dissolved in 1% *n*-octyl- β -D-glucopyranoside. Spectra were obtained before and after acid denaturation by addition of 1/10 volume 100 mM sulphuric acid.

Transducin activation assay. As a monitor for Gt activation, changes in intrinsic fluorescence intensity of the G α -subunit upon exchange of GDP to GTP γ S were quantified^{41–43}. All measurements were carried out using a SPEX fluorolog II spectrofluorometer equipped with a 450 W xenon arc lamp. For all activation measurements, settings were $\lambda_{\text{ex}} = 300$ nm and $\lambda_{\text{em}} = 345$ nm with an integration time of 1 s. Gt activation rates were measured with 2.5 nM opsin, 6 μ M GTP γ S, 2 mM DTT, and 0.006% (w/v) *n*-dodecyl- β -D-maltopyranoside in a final volume of 650 μ l (10 mm \times 4 mm cuvette with stirring bar). All samples were equilibrated at 20 °C for 4 min while being irradiated with yellow light (Schott GG 495 long-pass filter). Then reactions were triggered by addition of GTP γ S after recording basic fluorescence levels for 50 s. After recording the fluorescence change at the initial opsin concentration, activation of the whole Gt pool was achieved by adding 10 nM purified rhodopsin to the reaction. The concentration of intact opsin in every sample was calculated from the quantity of reconstituted rhodopsin by 500 nm absorption after addition of excess 11-*cis*-retinal to an aliquot of the sample measured. The 500 nm rhodopsin absorption was determined from difference spectra dark minus light-activated rhodopsin.

31. Herrmann, R. *et al.* Sequence of interactions in receptor-G protein coupling. *J. Biol. Chem.* **279**, 24283–24290 (2004).
32. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
33. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
34. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
35. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
36. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
37. Hoof, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272 (1996).
38. McDonald, I. K. & Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793 (1994).
39. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134 (1995).
40. DeLano, W. L. The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, California, USA, 2002).
41. Fahmy, K. & Sakmar, T. P. Regulation of the rhodopsin-transducin interaction by a highly conserved carboxylic acid group. *Biochemistry* **32**, 7229–7236 (1993).
42. Ernst, O. P., Bieri, C., Vogel, H. & Hofmann, K. P. Intrinsic biophysical monitors of transducin activation: fluorescence, UV-visible spectroscopy, light scattering, and evanescent field techniques. *Methods Enzymol.* **315**, 471–489 (2000).
43. Ernst, O. P., Gramse, V., Kolbe, M., Hofmann, K. P. & Heck, M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. *Proc. Natl Acad. Sci. USA* **104**, 10859–10864 (2007).

Aberrant chromosome morphology in human cells defective for Holliday junction resolution

Thomas Wechsler^{1†}, Scott Newman² & Stephen C. West¹

In somatic cells, Holliday junctions can be formed between sister chromatids during the recombinational repair of DNA breaks or after replication fork demise. A variety of processes act upon Holliday junctions to remove them from DNA, in events that are critical for proper chromosome segregation. In human cells, the BLM protein, inactivated in individuals with Bloom's syndrome, acts in combination with topoisomerase III α , RMI1 and RMI2 (BTR complex) to promote the dissolution of double Holliday junctions^{1,2}. Cells defective for BLM exhibit elevated levels of sister chromatid exchanges (SCEs) and patients with Bloom's syndrome develop a broad spectrum of early-onset cancers caused by chromosome instability³. MUS81–EME1 (refs 4–7), SLX1–SLX4 (refs 8–11) and GEN1 (refs 12, 13) also process Holliday junctions but, in contrast to the BTR complex, do so by endonucleolytic cleavage. Here we deplete these nucleases from Bloom's syndrome cells to analyse human cells compromised for the known Holliday junction dissolution/resolution pathways. We show that depletion of MUS81 and GEN1, or SLX4 and GEN1, from Bloom's syndrome cells results in severe chromosome abnormalities, such that sister chromatids remain interlinked in a side-by-side arrangement and the chromosomes are elongated and segmented. Our results indicate that normally replicating human cells require Holliday junction processing activities to prevent sister chromatid entanglements and thereby ensure accurate chromosome condensation. This phenotype was not apparent when both MUS81 and SLX4 were depleted from Bloom's syndrome cells, suggesting that GEN1 can compensate for their absence. Additionally, we show that depletion of MUS81 or SLX4 reduces the high frequency of SCEs in Bloom's syndrome cells, indicating that MUS81 and SLX4 promote SCE formation, in events that may ultimately drive the chromosome instabilities that underpin early-onset cancers associated with Bloom's syndrome.

Our current understanding of the way in which Holliday junctions are processed in somatic cells suggests the three potential pathways illustrated in Supplementary Fig. 1. These include the dissolution of double Holliday junctions by BLM–TOPIII α –RMI1–RMI2 (BTR), which suppresses crossover formation between sister chromatids¹, and the nucleolytic resolution of Holliday junctions by MUS81–EME1 (ref. 7) or GEN1 (ref. 12) that can lead to crossover or non-crossover products depending on the orientation of Holliday junction cleavage. Recently, it was shown that SLX4, a component of the SLX1–SLX4 nuclease complex that can also cleave Holliday junctions, associates with MUS81–EME1 and may provide a 'scaffold' function for several nuclease activities^{8–11}.

The relative contribution of each Holliday junction processing pathway is currently unknown. However, given that intact Holliday junctions are a relatively poor substrate for MUS81–EME1 (refs 4, 5), it is likely that the BTR complex provides the primary mechanism for the resolution of double Holliday junctions in human somatic cells at S phase. A further possibility is that resolution events mediated by MUS81–EME1, SLX1–SLX4 and/or GEN1 could substitute for the loss

of BTR activity in Bloom's syndrome cells, either by cleaving the double Holliday junctions or other recombination intermediate structures (such as nicked Holliday junctions^{14,15}), and thereby contribute to their viability.

Because nucleolytic cleavage mechanisms may be responsible for the elevated frequency of SCEs observed in Bloom's syndrome cells, we analysed SCE formation in metaphase spreads from the SV40-transformed Bloom's syndrome cell line GM08505 after short interfering RNA (siRNA)-mediated depletion of MUS81, SLX4 or GEN1. In all cases, efficient gene silencing was achieved, as measured by western blotting or quantitative reverse transcription PCR (RT–PCR) (Supplementary Fig. 2). Because depletion of SLX4 also decreases the levels of SLX1 (ref. 10), as the stabilities of SLX1 and SLX4 are interdependent, the SLX4 depletion should be viewed as an SLX1–SLX4 depletion. Depletion of SLX4 does not affect the levels of MUS81 or EME1 (ref. 10), or GEN1 (data not shown). We found that siRNA treatment against MUS81 or SLX4, but not GEN1, significantly reduced the frequency of SCEs (Fig. 1a, b), as well as the formation of harlequin chromosomes (that is, chromosomes exhibiting more than five SCEs) (Supplementary Fig. 3).

Although the SCE frequency in cells depleted for both MUS81 and GEN1 did not appear significantly different from MUS81-depleted cells, at least in metaphases that could be easily scored, many metaphase chromosomes looked abnormal after treatment with these siRNAs. Because these metaphases could not be scored for SCEs, it is possible that our scoring was biased towards those with only mild GEN1 and/or MUS81 depletion. We also observed decreased cell viability after dual siRNA treatment against GEN1 and MUS81, MUS81 and SLX4, or GEN1 and SLX4 (Fig. 1c), indicating that loss of multiple Holliday junction processing pathways can lead to cell death even in the absence of exogenous DNA damage. Depletion of GEN1 with SLX4, or MUS81 with SLX4, had a greater impact upon cell viability than GEN1 and MUS81 siRNA treatment, suggesting that SLX4 might have a broader role than either of the other nucleases.

When the chromosome aberrations seen in the metaphase spreads prepared from GEN1- and MUS81-depleted cells were analysed, we observed a high percentage of cells in which the chromosomes were elongated and segmented (Fig. 2a, compare left and right panels with enlargements below). Indeed, most chromosomes exhibited a 'beads-on-a-string' morphology. Similar results were obtained when the untransformed Bloom's syndrome fibroblast line GM01492 was treated with MUS81 and GEN1 siRNAs (Fig. 2b and Supplementary Fig. 4). Careful analysis of these spreads revealed the occasional appearance of single chromosomes that had an extreme defect in chromosome condensation (Fig. 2b, arrows).

These observations, with measurements showing that the total number of chromosomes in these metaphases was comparable to control spreads (data not shown), indicated that the beads-on-a-string morphology might result from an aberrant chromosome condensation defect rather than break-induced chromosome rearrangements. To confirm this, we performed whole chromosome painting on chromosomes 4, 8

¹London Research Institute, Cancer Research UK, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK. ²Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK. [†]Present address: Stanford University School of Medicine, Clark Center, 318 Campus Drive, Stanford, California 94305, USA.

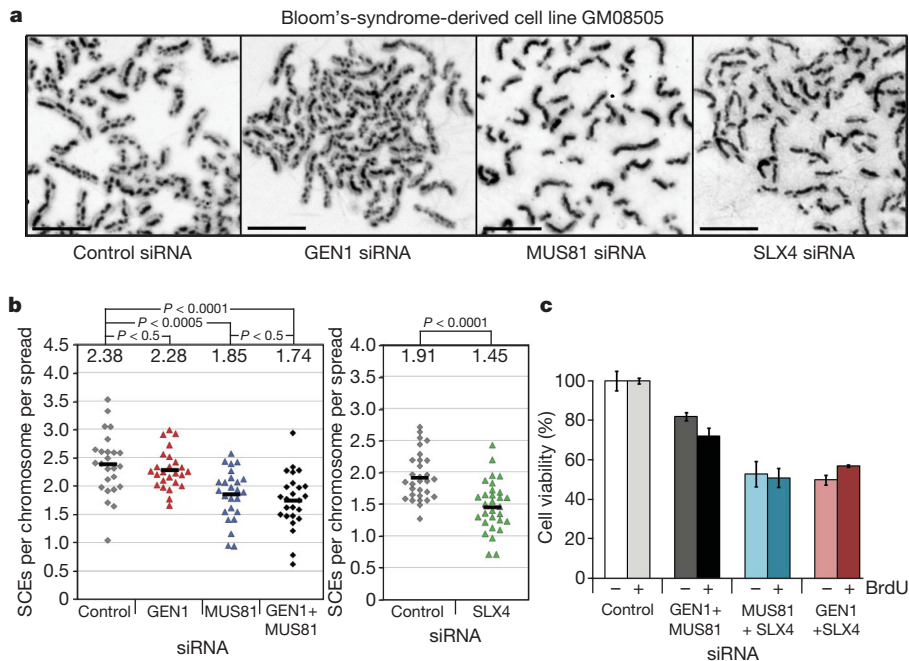


Figure 1 | Contribution of GEN1, MUS81 and SLX4 to SCE frequency in Bloom's syndrome cells. **a**, Representative images of metaphase spreads prepared from BLM-deficient GM08505 cells treated with the indicated siRNAs. Scale bars, 10 μ m. **b**, Quantification of SCE frequency after siRNA treatment. Each data point represents a single cell/metaphase that was scored

and X in GM01492 Bloom's syndrome cells after GEN1 and MUS81 siRNA targeting. We found that segmentation occurred within intact chromosomes and was not due to translocation events (Fig. 2c). Indeed, although some chromosomal regions looked compact, others were more

blind for SCEs per chromosome per spread (for each condition 25 cells, more than 1,700 chromosomes, were analysed). *P* values were determined using a two-tailed *t*-test. **c**, Relative cell viability measured 96 h after siRNA treatment that targeted two genes as indicated. Analyses were performed with and without BrdU treatment. Error bars, s.d.

elongated and appeared as though held together by a thread of (non-staining) DNA. Our interpretation of the extended and segmented chromosome structure is that regions showing normal condensation are linked to uncondensed regions of DNA. Generally, the uncondensed

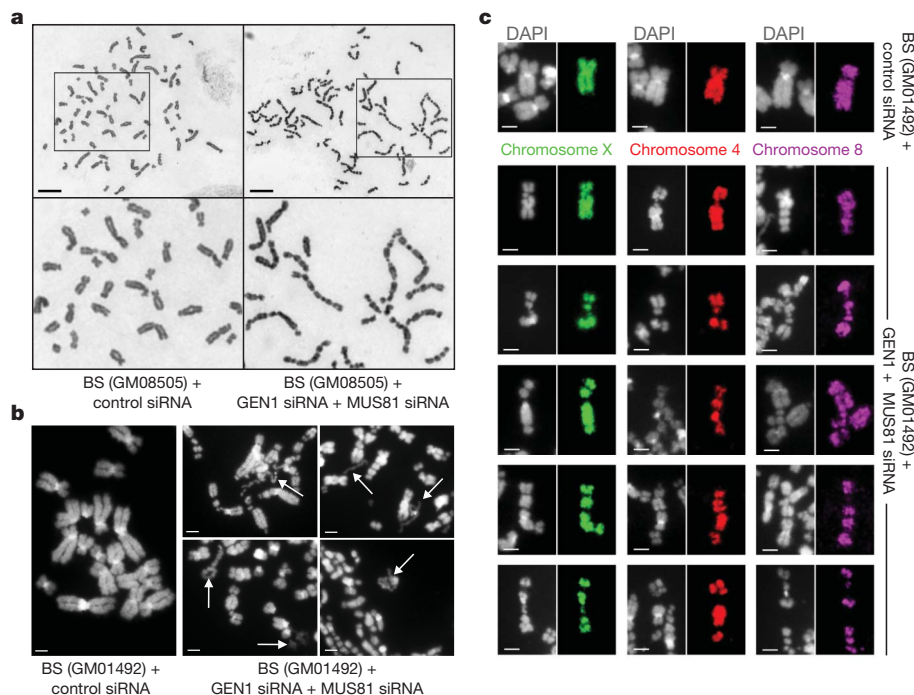


Figure 2 | Chromosome abnormalities in Bloom's syndrome (BS) cells after siRNA-mediated depletion of GEN1 and MUS81. **a**, Giemsa-stained metaphase spreads of GM08505 BLM-defective cells treated with control siRNA or siRNAs against both GEN1 and MUS81. The upper panels show whole chromosome spreads whereas the lower panels show the indicated sections at higher magnification. Scale bars, 10 μ m. **b**, 4',6-Diamidino-2-phenylindole (DAPI)-stained metaphase spreads of GM01492 BLM-defective

cells treated with control siRNA or siRNAs against both GEN1 and MUS81 as indicated. White arrows indicate chromosomes with extreme condensation defects. Scale bars, 2 μ m. **c**, Metaphase spreads from the experiment described in **b** were stained with whole chromosome paints specific for chromosome 4, 8 or X, as indicated. The selected chromosomes are representative of those from a total of nine metaphase spreads. Scale bars, 2 μ m.

regions were observed at equivalent positions on each of the two sister chromatids, which themselves remained tightly associated along their entire length. One possibility is that the loss of Holliday junction processing activity results, either directly or indirectly, in sister chromatid entanglements that prohibit normal chromosome condensation.

To quantify the observed segmentation phenotype, we counted chromosomes with more than three segments ($S > 3$) and scored metaphases with two or more segmented chromosomes as ' $S > 3$ positive' (Fig. 3a). Using this scoring method, a baseline of 2% was observed with the control GM08505 Bloom's syndrome cells, whereas cells depleted for MUS81 or GEN1 contained 6% and 8% of $S > 3$ chromosomes, respectively (Fig. 3b). In contrast, 56% of the Bloom's syndrome metaphases depleted for both GEN1 and MUS81 exhibited two or more $S > 3$ chromosomes (Fig. 3b, c). Moreover, in 13% of these cells, the $S > 3$ phenotype was so severe that virtually all chromosomes in the spread were affected (Fig. 2a). To enable the most simple quantification, these severely affected cells were designated '>9'. The abnormal phenotype was not influenced by the presence or absence of BrdU (an agent used in the earlier SCE analyses) (Supplementary Fig. 5). When the same scoring method was applied to the Bloom's syndrome cell line GM01492 after GEN1- and MUS81-depletion, we found that 46% of the cells (compared with 7% in control cells) showed an $S > 3$ phenotype, ruling out any cell-line-specific effects (Supplementary Fig. 4).

Importantly, the chromosome abnormalities observed in Bloom's syndrome cells depleted for both GEN1 and MUS81 were suppressed by exogenous expression of BLM protein, as shown by comparison of the BLM-defective cell line PSNG13 with its isogenic BLM-complemented cell line PSNF5 after GEN1 and MUS81 siRNA treatment (Fig. 3d–f and Supplementary Fig. 6). We also failed to observe

an increase in $S > 3$ chromosomes in the BLM-proficient cell line U2OS, despite a high depletion efficiency of GEN1 and MUS81 (Supplementary Fig. 7). These results show that BLM is critical for the maintenance of genome stability, and that loss of Holliday junction processing activity caused by disruption of BLM, MUS81 and GEN1 leads to aberrant chromosome morphology.

In yeast, it has been shown that *yen1 mus81* double mutants (Yen1 is the yeast orthologue of GEN1) are considerably more sensitive to DNA damage than the *mus81* single mutant^{16,17} and that Mus81 and Yen1 can promote crossover formation during mitotic recombination¹⁸. These studies indicate that recombination intermediates normally resolved by Mus81 can also serve as substrates for Yen1. In mammalian cells, however, we currently have little information relating to the interplay between GEN1 and MUS81, or with SLX4 with which MUS81 interacts^{8–11}. Therefore, to gain our first insights into the genetic interactions between these proteins, we depleted combinations of either MUS81 + SLX4, GEN1 + SLX4 or, as before, GEN1 + MUS81, and measured the extent of chromosome aberrations. We found that depletion of GEN1 and SLX4 in Bloom's syndrome cells resulted in an extremely severe phenotype, as measured by the formation of segmented chromosomes (Fig. 4). The next most severe combination was caused by depletion of GEN1 and MUS81. In contrast, the chromosome abnormalities observed after siRNA treatment against both MUS81 and SLX4 were considerably less severe. These results favour the view that SLX4 plays a broad role, not only in a nuclease complex with SLX1, but also as a scaffold for the cooperative actions of other nucleases.

The severe phenotype observed after GEN1 and SLX4 depletion allowed us to perform two further experiments. First, time-course analyses revealed that $S > 3$ chromosomes were detectable 36 h after

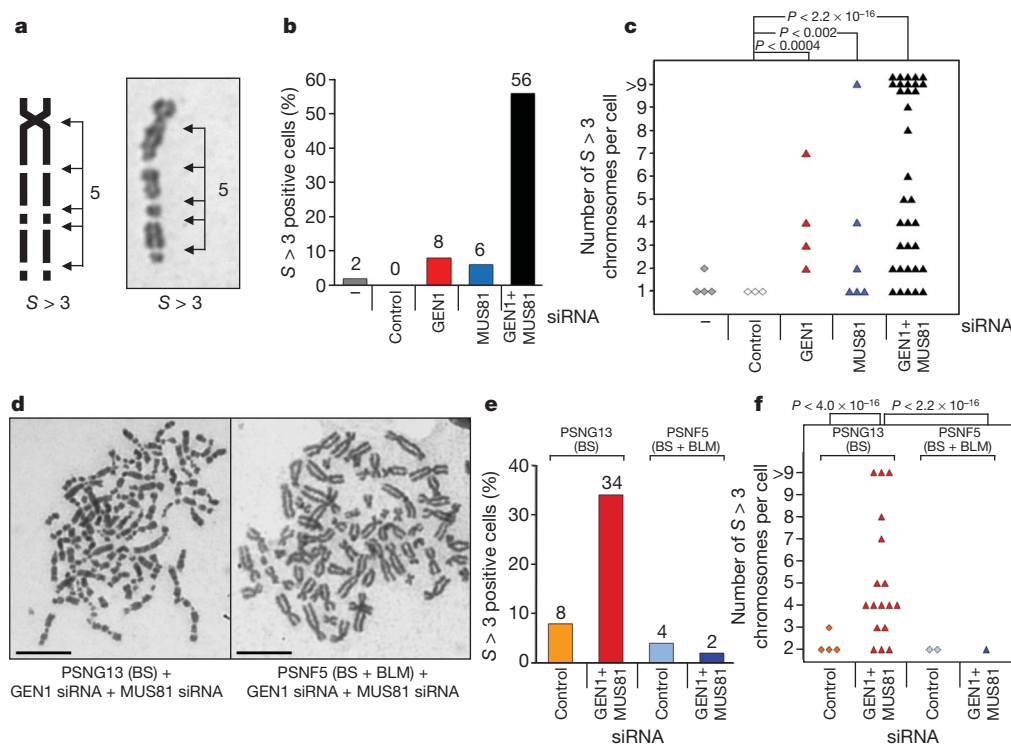


Figure 3 | Quantification of the chromosome segmentation phenotype observed in Bloom's syndrome cells after GEN1 and MUS81 depletion. **a**, Schematic illustration (left) and Giemsa staining (right) of an abnormal metaphase chromosome showing more than three indentations ($S > 3$). The definition $S > 3$ was used in all subsequent experiments. **b**, After the treatment of GM08505 Bloom's syndrome cells with the indicated siRNAs, metaphases ($n = 50$) were scored. The percentage of cells with at least two $S > 3$ chromosomes is shown. **c**, As **b** but to illustrate phenotypic severity, the number of $S > 3$ chromosomes in each metaphase ($n = 50$) was plotted in a

scatter graph. Only metaphases with at least one $S > 3$ chromosome are shown, and those with more than nine segmented chromosomes were termed '>9'. **d**, Complementation of Bloom's syndrome cells with BLM restores chromosome stability. BLM-deficient PSNG13 and the BLM-complemented PSNF5 cell lines were treated with siRNA against GEN1 and MUS81, and analysed for abnormal metaphases using Giemsa staining. Scale bars, 10 μ m. **e**, **f**, Metaphases ($n = 50$) were analysed after control or GEN1 + MUS81 siRNA treatment of PSNG13 and PSNF5 cells. Quantifications were performed as in Fig. 3b.

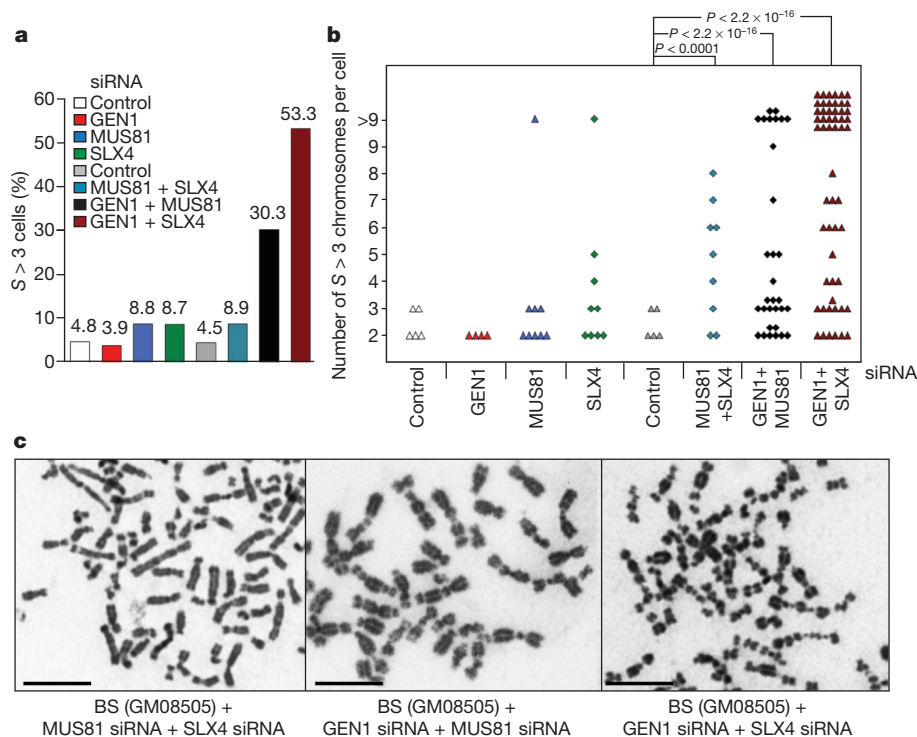


Figure 4 | Synthetic interactions between GEN1, MUS81 and SLX4.

GM08505 Bloom's syndrome cells were depleted for GEN1, MUS81 or SLX4, or for the indicated pairs of proteins. Cells were collected 60 h after the second

transfection and metaphase cells ($n > 100$) were scored (a, b) and visualized (c) as in Fig. 3. Scale bars, 10 μ m.

siRNA treatment, and that the severity of the abnormal chromosome phenotype continued into the second mitotic division (Supplementary Fig. 8). After this point, high levels of cell inviability were observed (Fig. 1c). Second, we found that the chromosome indentations were effectively free of SMC2 protein, confirming that the unusual phenotype was indeed due to a defect in proper chromosome condensation (Supplementary Fig. 9).

These results begin to define the relative contributions of GEN1, MUS81 and SLX4 to Holliday junction processing in human mitotic cells already defective for BLM. Inactivation of MUS81, together with GEN1, effectively produces BLM-defective cells compromised for the known Holliday junction dissolution/resolution pathways. The resulting phenotype was the formation of highly segmented chromosomes with severe condensation defects. Similar results were observed by inactivation of GEN1 and SLX4, supporting the notion that SLX4 and MUS81 might work in the same pathway. Although this is the first time that such an aberrant chromosome morphology has been associated with defects in Holliday-junction processing, similar defects have been observed in human and mouse cells in response to ionizing radiation which causes a delay in DNA replication timing that, in turn, affects mitotic chromosome condensation¹⁹. Similarly, defects in DNA replication due to the mutation of key replication factors such as ORC2 also lead to segmented chromosome condensation morphology. Together, these studies indicate a potential link between DNA replication and the establishment of proper chromosome condensation, and it has been suggested that the timely completion of replication impacts upon the lateral condensation of a metaphase chromosome by helping to remove entanglements²⁰. In our study, we suggest that sister chromatid entanglements, caused instead by defects in Holliday junction processing pathways, might lead to a related aberrant condensation phenotype. The beads-on-a-string chromosome morphology seen here, and the absence of significant break-induced chromosome translocations, indicates that the regions of under-condensed chromatin retain fine DNA threads that link the normally condensed regions.

Moreover, the tight side-by-side alignment of sister chromatids is suggestive of the persistence of unresolved bridges that accumulate after disruption of the three cellular Holliday junction processing pathways.

It has been shown that BLM protein localizes to ultrafine bridges at anaphase, and that these sites represent unresolved entanglements that occur at fragile sites after replication stress^{21–23}. However, many ultrafine bridges derive from centromeric regions^{21,24}, whereas our work shows that the indentations corresponding to regions of under-condensed chromatin are distributed randomly along the length of each chromosome.

The aberrant chromosome morphology seen in the current experiments was suppressed by expression of the BLM protein, leading us to suggest that Holliday junction dissolution, a system that avoids SCEs, provides the primary mechanism for the processing of Holliday junctions in somatic cells. In the absence of this pathway, such as in cells derived from patients with Bloom's syndrome, elevated levels of SCEs are observed and we have shown that these can arise through the actions of MUS81 or SLX4. The precise role of SLX4 is currently unknown, as it may function either as a junction-specific nuclease with SLX1, or it may provide a scaffold for the cooperation of multiple nucleases within a multi-functional DNA processing complex. Loss of both BLM and MUS81 (or SLX4), however, did not result in a severe condensation phenotype, because mitotic cells possess a third pathway of Holliday junction resolution mediated by GEN1, for which we present the first functional evidence *in vivo*. It is likely that MUS81–EME1, SLX1–SLX4 and GEN1 can resolve Holliday junctions that persist in BLM-defective cells, although the bi-directional nature of their cleavage mechanism will, in contrast to the BTR complex, produce SCEs. We therefore suggest that the nucleolytic processing pathways provide additional mechanisms of resolution that can act upon intermediates that escape the attention of the BTR complex, thereby allowing chromosome segregation. Use of these alternatives may, however, come at a price because Bloom's syndrome cells exhibit genomic instability and patients suffer a broad spectrum of early-onset cancers.

METHODS SUMMARY

The human osteosarcoma cell line U2OS, SV40-transformed Bloom's syndrome fibroblasts GM08505 (ref. 25) and untransformed Bloom's syndrome fibroblasts GM01492 (ref. 26) were provided by CRUK Cell Services (Clare Hall). The cell line PSNF5 constitutively expresses BLM protein, whereas the isogenic control PSNG13 contains pcDNA3 vector DNA²⁷. Cells were transfected twice with siRNA within 24 h. For a single gene, 3×10^5 cells were transfected with 400–500 pmol of siRNA (Dharmacon) using 10 μ l of Lipofectamine RNAiMAX (Invitrogen). For double-gene targeting, up to 800 pmol of siRNA was mixed and transfected with 20 μ l Lipofectamine. Control non-targeting siRNAs were used at equivalent concentrations. The efficiency of each siRNA treatment was monitored by western blotting 60 h after the first, or 36 h after the second, transfection, or by quantitative RT-PCR 48 h after the first treatment. SCEs and metaphase chromosomes were monitored as described^{28,29}. The SCE technique relies upon the differential staining of the sister chromatids after replication in the presence of BrdU, followed by staining with Hoechst dye and Giemsa. Cell viability assays were performed 12 h after the second transfection with various siRNA combinations (MUS81 + SLX4; GEN1 + MUS81; GEN1 + SLX4). The cells were then seeded in equal numbers (2,000 per well) in a 96-well plate (in triplicates), grown for a further 84 h and tested using the Cell-Titre Glo assay system (Promega). Chromosome painting was performed as described³⁰.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 July 2010; accepted 5 January 2011.

Published online 13 March 2011.

- Wu, L. & Hickson, I. D. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870–874 (2003).
- Mankouri, H. W. & Hickson, I. D. The RecQ helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific 'dissolvasome'? *Trends Biochem. Sci.* **32**, 538–546 (2007).
- Bachrati, C. Z. & Hickson, I. D. RecQ helicases, suppressors of tumorigenesis and premature aging. *Biochem. J.* **374**, 577–606 (2003).
- Chen, X. B. *et al.* Human MUS81-associated endonuclease cleaves Holliday junctions *in vitro*. *Mol. Cell* **8**, 1117–1127 (2001).
- Ciccia, A., Constantinou, A. & West, S. C. Identification and characterization of the human MUS81/EME1 endonuclease. *J. Biol. Chem.* **278**, 25172–25178 (2003).
- Ciccia, A., McDonald, N. & West, S. C. Structural and functional relationships of the XPF/MUS81 family of proteins. *Annu. Rev. Biochem.* **77**, 259–287 (2008).
- Taylor, E. R. & McGowan, C. H. Cleavage mechanism of human MUS81–EME1 acting on Holliday-junction structures. *Proc. Natl Acad. Sci. USA* **105**, 3757–3762 (2008).
- Andersen, S. L. *et al.* *Drosophila* MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. *Mol. Cell* **35**, 128–135 (2009).
- Fekairi, S. *et al.* Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* **138**, 78–89 (2009).
- Munoz, I. M. *et al.* Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Mol. Cell* **35**, 116–127 (2009).
- Svendsen, J. M. *et al.* Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* **138**, 63–77 (2009).
- Ip, S. C. Y. *et al.* Identification of Holliday junction resolvases from humans and yeast. *Nature* **456**, 357–361 (2008).
- Rass, U. *et al.* Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev.* **24**, 1559–1569 (2010).
- Osman, F., Dixon, J., Doe, C. L. & Whitby, M. C. Generating crossovers by resolution of nicked Holliday junctions: a role of Mus81–Eme1 in meiosis. *Mol. Cell* **12**, 761–774 (2003).
- Gaillard, P.-H. L., Noguchi, E., Shanahan, P. & Russell, P. The endogenous Mus81–Eme1 complex resolves Holliday junctions by a nick and couter-nick mechanism. *Mol. Cell* **12**, 747–759 (2003).
- Blanco, M. G., Matos, J., Rass, U., Ip, S. C. Y. & West, S. C. Functional overlap between the structure-specific nucleases Yen1 and Mus81–Mms4 for DNA damage repair in *S. cerevisiae*. *DNA Repair (Amst.)* **9**, 394–402 (2010).
- Tay, Y. D. & Wu, L. Overlapping roles for Yen1 and Mus81 in cellular Holliday junction processing. *J. Biol. Chem.* **285**, 11427–11432 (2010).
- Ho, C. K., Mazón, G., Lam, A. F. & Symington, L. S. Mus81 and Yen1 promote reciprocal exchange during mitotic recombination to maintain genome integrity in budding yeast. *Mol. Cell* **40**, 988–1000 (2011).
- Breger, K. S., Smith, L., Turker, M. S. & Thayer, M. J. Ionizing radiation induces frequent translocations with delayed replication and condensation. *Cancer Res.* **64**, 8231–8238 (2004).
- Hearst, J., Kauffman, L. & McClain, W. A simple mechanism for the avoidance of entanglement during chromosome replication. *Trends Genet.* **14**, 244–247 (1998).
- Chan, K. L., North, P. S. & Hickson, I. D. BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *EMBO J.* **26**, 3397–3409 (2007).
- Chan, K. L., Palmai-Pallag, T., Ying, S. M. & Hickson, I. D. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nature Cell Biol.* **11**, 753–760 (2009).
- Naim, V. & Rosselli, F. The FANCD pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities. *Nature Cell Biol.* **11**, 761–768 (2009).
- Baumann, C., Korner, R., Hofmann, K. & Nigg, E. A. PICH, a centromere-associated SNF2 family ATPase, is regulated by PIK1 and required for the spindle checkpoint. *Cell* **128**, 101–114 (2007).
- Wu, L., Davies, S. L., Levitt, N. C. & Hickson, I. D. Potential role for the BLM helicase in recombinational repair via a conserved interaction with RAD51. *J. Biol. Chem.* **276**, 19375–19381 (2001).
- Ellis, N. A., Proytcheva, M., Sanz, M. M., Ye, T.-Z. & German, J. Transfection of BLM into cultured Bloom syndrome cells reduced the sister-chromatid exchange rate toward normal. *Am. J. Hum. Genet.* **65**, 1368–1374 (1999).
- Gaymes, T. J. *et al.* Increased error-prone non homologous DNA end-joining – a proposed mechanism of chromosomal instability in Bloom's syndrome. *Oncogene* **21**, 2525–2533 (2002).
- Bender, C. F. *et al.* Cancer predisposition and hematopoietic failure in *Rad50*^(S/S) mice. *Genes Dev.* **16**, 2237–2251 (2002).
- Bayani, J. & Squire, J. A. Sister chromatid exchange. *Curr. Protoc. Cell Biol.* **22**, 7, (2005).
- Alsop, A. E., Teschendorff, A. E. & Edwards, P. A. W. Distribution of breakpoints on chromosome 18 in breast, colorectal, and pancreatic carcinoma cell lines. *Cancer Genet. Cytogenet.* **164**, 97–109 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank I. Hickson for providing the Bloom's syndrome cell lines and advice, P. Edwards for help and providing facilities for chromosome painting, S. Horswell for the statistical analysis, S. Ip for the GEN1 antibody, M.G. Blanco for assistance with SCE scoring and our laboratory colleagues for their encouragement and suggestions. We further thank K. Cimprich, the Cimprich laboratory members, and C. Wang, W. Johnson and A. Straight. This work was supported by Cancer Research UK, the Louis-Jeantet Foundation, the European Research Council, the Swiss Bridge Foundation and the Breast Cancer Campaign. S.N. was supported by a studentship from the UK Medical Research Council.

Author Contributions T.W. and S.C.W. designed the project that was undertaken entirely by T.W. Expertise for the chromosome paints was provided by S.N. The manuscript was written by S.C.W. with help from T.W.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.C.W. (stephen.west@cancer.org.uk).

METHODS

Antibodies. Affinity-purified anti-GEN1 rabbit polyclonal antibody was raised against a carboxy (C)-terminal peptide (CLDSPLPLRQLKLRQST) corresponding to GEN1^{890–908}. Mouse monoclonal antibodies against MUS81 (2G10/3) and RAD51 (14B4) were purchased from Abcam, and rabbit polyclonal antibody against SMC2 (A300-058A) was from Bethyl Labs.

siRNA transfections, quantitative RT-PCR and western blotting. For siRNA transfections, 3×10^5 cells were seeded in 60 mm cell culture dishes 8 h before transfection. In general, the cells were transfected twice within 24 h and the cell culture medium was changed after the first transfection. For depletion of a single protein, 3×10^5 cells were transfected with 400–500 pmol of siRNA using 10 μ l of Lipofectamine RNAiMAX (Invitrogen). For double targeting, up to 800 pmol of siRNA was mixed and transfected with 20 μ l Lipofectamine RNAiMAX; ON TARGET-plus siRNAs specific for each gene and control non-targeting siRNAs were purchased from Dharmacon and were used at the same concentrations. The efficiency of each siRNA treatment was monitored by western blotting 60 h after the first transfection, or 36 h after the second, or by quantitative RT-PCR 48 h after the first transfection. The siRNA sequences, indicated 5' to 3', were as follows. Control siRNA: ON-TARGETplus non-targeting siRNA #3. GEN1/FLJ40869 oligonucleotides 1–4: (1) GCGUAAUCUUGGUGGAAA; (2) UCUAAGACCUUUGGCUAUA; (3) UAUGCAAACACUCGGA; (4) GCCCUAAGAUACAUAUUA. SLX4 (oligonucleotides 1–2)^{9–11}: (1) AAACGUGAAGAAGCA GAAUU; (2) CGGCAUUUGAGUCGAGGUGAA. MUS81 (oligonucleotides 1–2): (1) CAGCCUGGUGGAUCGAUA; (2) CAUUAAGUGUGGCGUCUA.

Owing to the lack of specificity of SLX4 antibodies, the efficiency of depletion of SLX4 was determined by quantitative RT-PCR. For this, total RNA was isolated using the RNeasy kit (Qiagen) according to the manufacturer's instructions with the following modifications: cell lysis was achieved using QIAshredder columns (Qiagen) and DNase I treatment was performed before total RNA was eluted from the RNeasy column. Then 1 μ g of total RNA was reverse transcribed using a TaqMan Reverse Transcription Kit (Applied Biosystems) in a 20- μ l reaction and used for three quantitative RT-PCR reactions in a 96-well format using the EXPRESS SYBR GreenER qPCR SuperMix with Premixed ROX (Invitrogen) and a Real-Time PCR System (Applied Biosystems). The ribosomal protein L23 mRNA was used as a standard. Depletion of SLX4 (or SLX1) by siRNA treatment results in a loss of both SLX1 and SLX4, as their stabilities are interdependent¹⁰. Use of siRNAs against either is therefore considered to be equivalent to SLX1–SLX4 depletion.

Primers for quantitative RT-PCR of GEN1: forward, CCACATGACTATG AATACTGCTGTCCTT; backward, TGGGAATCCCTCACACAGCAAGC.

Primers for quantitative RT-PCR of SLX4: forward, CCTGGAGAAAA GGGTTTGT; backward, AGCTTCATCCAAGCACCTGT.

Primers for quantitative RT-PCR of L23: forward, TTCCTGGTCCACA ACGTCAAG; backward, TTGTGAAGCGATCTCGGCA.

Cell viability assays. Cells were collected 12 h after the second siRNA transfection with various siRNA combinations (MUS81 + SLX4; GEN1 + MUS81; GEN1 + SLX4). They were then seeded in equal numbers (2,000 per well) in 96-well plates in triplicates (with or without 66 μ M BrdU) and incubated for 84 h. Cell viability was determined using the Cell-Titre Glo assay (Promega).

Analysis of metaphase chromosomes. For the SCE assay, 3×10^5 cells were seeded in 60-mm plates before siRNA transfection. After 8 h, the cells were transfected with siRNA, grown for a further 24 h and then transfected again. After 18 h growth, BrdU (100 μ M) was added and the cells were grown for a further 60–72 h. Metaphase chromosomes were prepared and assayed for SCEs by a modification of published procedures^{28,29}. Briefly, cells were incubated for 1 h with 0.2 μ g ml^{−1} colcemid and metaphase cells were harvested by mitotic shake-off. The cells were then swollen in 75 mM KCl for 20 min, fixed with methanol:acetic acid (3:1) and spread. After treatment with Hoechst 33258 and ultraviolet treatment, images were acquired using a Zeiss Axio Imager M1 microscope using a Plan-Neofluar $\times 60$, 0.4 numerical aperture oil objective lens, and captured using an ORCA-ER camera (Hamamatsu) controlled by Volocity 4.3.2 software (Improvision). At least 25 images were taken randomly from each condition. The files were renamed and each image (at least 1,700 chromosomes per condition) was scored blind to determine both the number of SCEs per chromosome and the number of harlequin chromosomes per metaphase spread.

To visualize the segmented chromosome phenotype, the same procedure was performed except that BrdU was omitted and the cells were stained in 7% Giemsa for 7 min immediately after spreading and drying. The $S > 3$ phenotype was scored, and spreads with two or more chromosomes exhibiting the $S > 3$ phenotype were considered positive. For the most severely affected metaphases, the count was stopped at 10 and designated ' >9 '.

SMC2 staining. Metaphase cells were collected as above, swollen in $1 \times$ PME (5 mM Pipes/NaOH pH 7.2, 5 mM NaCl, 5 mM MgCl₂, 1 mM EGTA), resuspended in lysis buffer ($1 \times$ PME supplemented with protease inhibitors, 0.1% Triton-X, 1 mM ATP, 0.2 mM spermine, 0.5 mM spermidine and 10 μ g ml^{−1} cytochalasin B) and lysed on ice using a dounce homogenizer. Lysates were layered onto sucrose gradients (30, 40, 50 and 60% sucrose in $1 \times$ PME) and spun for 30 min at 2,000g. The chromosomes were taken from the 40/50% and 50/60% interfaces, and fixed in 3.7% formaldehyde and then centrifuged (20 min at 4,000g) through a 40% glycerol cushion onto polylysine-treated coverslips. Then chromosomes were stained with anti-SMC2 antibody (1:300) and Alexa Fluor 488-coupled goat anti-rabbit secondary antibody (1:1,000) and mounted in Vectashield containing DAPI.

Statistical analysis. Sister chromatid exchange data was subjected to a Student's two-tailed *t*-test. For $S > 3$ data, the counts were analysed by analysis of deviance using the generalized linear model fitting function glm() with a Poisson approximation to the multinomial, followed by analysis of deviance using the anova() function with the χ^2 test, both performed in R 2.10.1 (for details see <http://www.R-project.org>).

Tumour evolution inferred by single-cell sequencing

Nicholas Navin^{1,2}, Jude Kendall¹, Jennifer Troge¹, Peter Andrews¹, Linda Rodgers¹, Jeanne McIndoo¹, Kerry Cook¹, Asya Stepansky¹, Dan Levy¹, Diane Esposito¹, Lakshmi Muthuswamy³, Alex Krasnitz¹, W. Richard McCombie¹, James Hicks¹ & Michael Wigler¹

Genomic analysis provides insights into the role of copy number variation in disease, but most methods are not designed to resolve mixed populations of cells. In tumours, where genetic heterogeneity is common^{1–3}, very important information may be lost that would be useful for reconstructing evolutionary history. Here we show that with flow-sorted nuclei, whole genome amplification and next generation sequencing we can accurately quantify genomic copy number within an individual nucleus. We apply single-nucleus sequencing to investigate tumour population structure and evolution in two human breast cancer cases. Analysis of 100 single cells from a polygenomic tumour revealed three distinct clonal subpopulations that probably represent sequential clonal expansions. Additional analysis of 100 single cells from a monogenomic primary tumour and its liver metastasis indicated that a single clonal expansion formed the primary tumour and seeded the metastasis. In both primary tumours, we also identified an unexpectedly abundant subpopulation of genetically diverse ‘pseudodiploid’ cells that do not travel to the metastatic site. In contrast to gradual models of tumour

progression, our data indicate that tumours grow by punctuated clonal expansions with few persistent intermediates.

In single-nucleus sequencing (SNS), we isolate nuclei by flow-sorting and amplify DNA using whole genome amplification (WGA) for massively parallel sequencing (Supplementary Fig. 1). We achieve low coverage (~6%) of the genome of a single cell, sufficient to quantify copy number from sequence read depth. Several features of our data analysis were designed for SNS and differ from previous methods^{4–6} for measuring copy number from sequencing data. In contrast to using fixed intervals to calculate copy number, we use variable length bins but with uniform expected unique counts, which correct for biases that have been reported^{7–9} in WGA (Supplementary Fig. 2; see Methods). For each single cell, we typically achieve a mean read density of 138 per bin (standard error of the mean (s.e.m.) ± 5.55 , $n = 200$). Over-replicated loci called ‘pileups’, which have been previously reported in WGA^{10–12}, do occur in our data but not at recurrent locations in different cells (Supplementary Fig. 3). Pileups are sufficiently randomly distributed and sparse so as not to affect counting at the resolution we

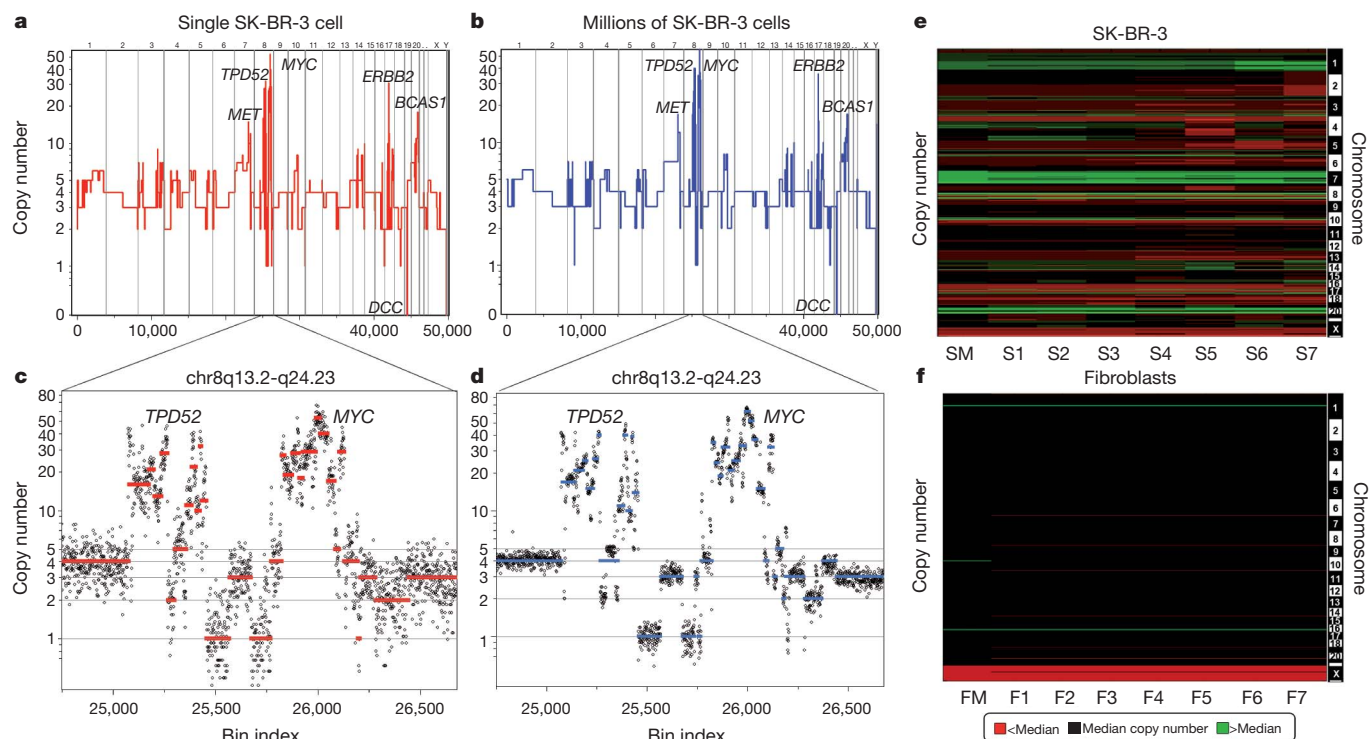


Figure 1 | Comparison of SK-BR-3 single cells to millions. **a, b**, The integer copy number profile for a single SK-BR-3 cell is shown (**a**) compared to a sequence count profile using millions of cells (**b**). **c, d**, A region on chromosome 8q13.2-q24.23 is plotted showing the integer copy number profile (in red or blue) and a ratio of raw bin counts in grey for a single cell (**c**), and a million cells

(**d**). **e**, A heatmap of SK-BR-3 copy number profiles comparing a million-cell sample (SM) to seven single cells (S1–S7). **f**, A heatmap of SK-N normal fibroblast profiles comparing a million-cell sample (FM) to seven single cells (F1–F7).

¹Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. ²Department of Genetics, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ³Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada.

have chosen (54 kb). Assuming that single cells will have discrete copy number states, we segment the variable bins and calculate integer copy number profiles (Supplementary Fig. 4; see Methods).

To validate our method, we compared the sequence counting profile of DNA from a single SK-BR-3 cell (Fig. 1a) with DNA from one million cells (Fig. 1b). The major amplifications (*MET*, *TPD52*, *ERBB2*, *BCAS1*) and deletions (*DCC*) are detected in both profiles, as are much more abundant but less marked small changes in copy number. To demonstrate how reproducible small differences are, we assessed data for a complex region on chromosome 8q13.2-q24.23 that contains more than thirty segments with differing copy number. These data were reproducible in both a single-cell (Fig. 1c) and a million-cell sample (Fig. 1d). We also compared the sequence read profiles from several single cells and from a million cells to each other and to the profile measured by microarray comparative genomic hybridization (CGH) from bulk DNA (Supplementary Fig. 5). In all instances the profiles showed very high ($r^2 > 0.85$) correlation. The reproducibility

and variation between single-cell copy number profiles was also investigated by comparing seven single cells from a culture of SK-BR-3 and seven from normal human fibroblasts. These data are shown as heat maps (Fig. 1e–f), which show that some genomic variation exists between cells. The diploid fibroblast cultures showed no random events; we observed only a few consistent events at levels expected for heritable copy number variations.

We selected next two high-grade (III), triple-negative (ER[−], PR[−], HER2[−]) ductal carcinomas (T10, T16P) and a paired metastatic liver carcinoma (T16M) to study tumour population structure and infer tumour evolution by single-cell analysis. T10 was selected to study primary tumour growth because it was previously shown¹³ to be genetically heterogeneous (polygenomic), and T16P was selected because it was classified as genetically homogeneous (monogenomic).

T10 was macrodissected into 12 sectors to preserve anatomical information, and nuclei were flow-sorted from six sectors (S1–S6) for SNS (Fig. 2a). Fluorescence-activated cell sorting (FACS) analysis

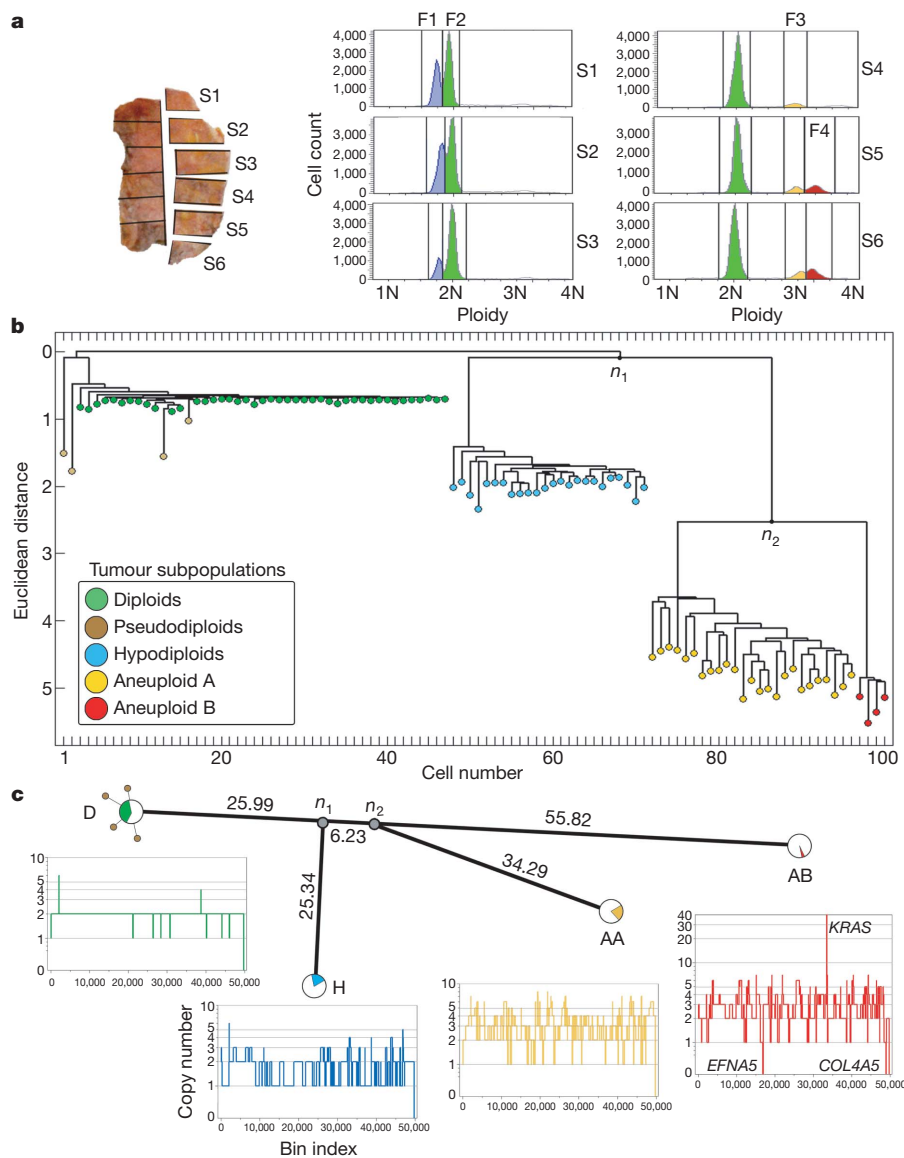


Figure 2 | Analysis of 100 single cells from a polygenomic breast tumour. **a**, T10 was macrodissected into 12 sectors, and nuclei were isolated from six sectors and flow-sorted by ploidy. FACS profiles show four distributions of ploidy (F1–F4), which were gated to isolate 100 single cells. **b**, Neighbour-joining tree of integer copy number profiles showing four major branches of

evolution. **c**, Phylogenetic tree of consensus profiles show the common ancestors and evolutionary distance between subpopulations. Integer copy number profiles from single cells are displayed below, and pie charts indicate the percentage of cells that constitute each subpopulation.

showed four major distributions of ploidy: a hypodiploid fraction (F1) exclusive to sectors 1–3; a diploid 2N fraction (F2) in all sectors; and two subtetraploid fractions (F3 and F4) in sectors 4–6. We selected 100 single cells from multiple sectors and ploidy fractions for sequencing and calculation of integer copy number profiles (Supplementary Table 1).

Breast tumours are typically mixtures of cancer cells with normal tissue, stroma and infiltrating leukocytes. By histopathology, T10 was assessed to contain 63% normal and 37% tumour cells and noted to be heavily infiltrated with leukocytes. Most of the diploid nuclei from F2 had flat genome profiles, characteristic of normal cells. Nearly two-thirds (31/47) of these diploid profiles showed narrow deletions in the T-cell receptor loci or one or more immunoglobulin variable region loci, consistent with infiltration by immunocytes (data not shown). Of the remaining sixteen nuclei from F2, twelve showed no discernable aberrations, but four nuclei showed aberrant profiles with diverse chromosome gains and losses. Each of these ‘pseudodiploid’ nuclei profiles seemed unrelated to the others or to those of the major tumour cell populations found in fractions F1, F3 and F4.

To determine population substructure we calculated pair-wise distances between the 100 integer copy number profiles, and built a tree using neighbour joining¹⁴ (Fig. 2b). The 100 profiles clustered into four subpopulations (D+P, H, AA and AB) regardless of their sector of origin. The D+P subpopulation contains predominantly flat diploid (D) profiles, but also pseudodiploid (P) cells that have diverged by varying degrees from the diploids. The three major ‘advanced’ tumour subpopulations (H, AA and AB) are highly clonal with complex genomic rearrangements, and together comprise slightly less than half the

cells of the tumour. These cells were isolated from the hypodiploid (F1) and two subtetraploid (F3 and F4) ploidy fractions, respectively. We had previously identified these subpopulations by profiling millions of cells by array CGH¹³, but we could not determine if they were composite mixtures of different tumour clones. By SNS we can now see that each subpopulation is composed of cells that share highly similar copy number profiles, probably representing three clonal expansions. Each subpopulation (H, AA and AB) is clearly related to the others by many shared genomic alterations, but they have also diverged and developed distinct attributes (for example, a massive 50-fold amplification of the *KRAS* oncogene in AB). The H cells display the characteristic ‘sawtooth’ pattern¹⁵ comprising broad chromosomal deletions (Fig. 2c). They are anatomically segregated in sectors S1–S3 of the tumour, whereas the AA and AB clones are intermixed and occupy sectors S4–S6.

To understand the relationship between subpopulations, we clustered profiles by chromosome breakpoints (which are directly related to the steps by which tumour cells diverge). We identified 657 copy number breakpoints and used them to build a phylogenetic tree, which closely resembles the structure of the neighbour-joining tree based on copy number (Supplementary Fig. 6). We also applied biclustering¹⁶ to construct a heat map of breakpoints, and ordered it on the basis of the copy number tree to show which breakpoints were common or divergent between the major subpopulations (Supplementary Fig. 7a). Although there is considerable variation within each subpopulation, no obvious further population substructure was evident. To estimate the common ancestors, we constructed a phylogenetic lineage using the consensus breakpoint patterns from the

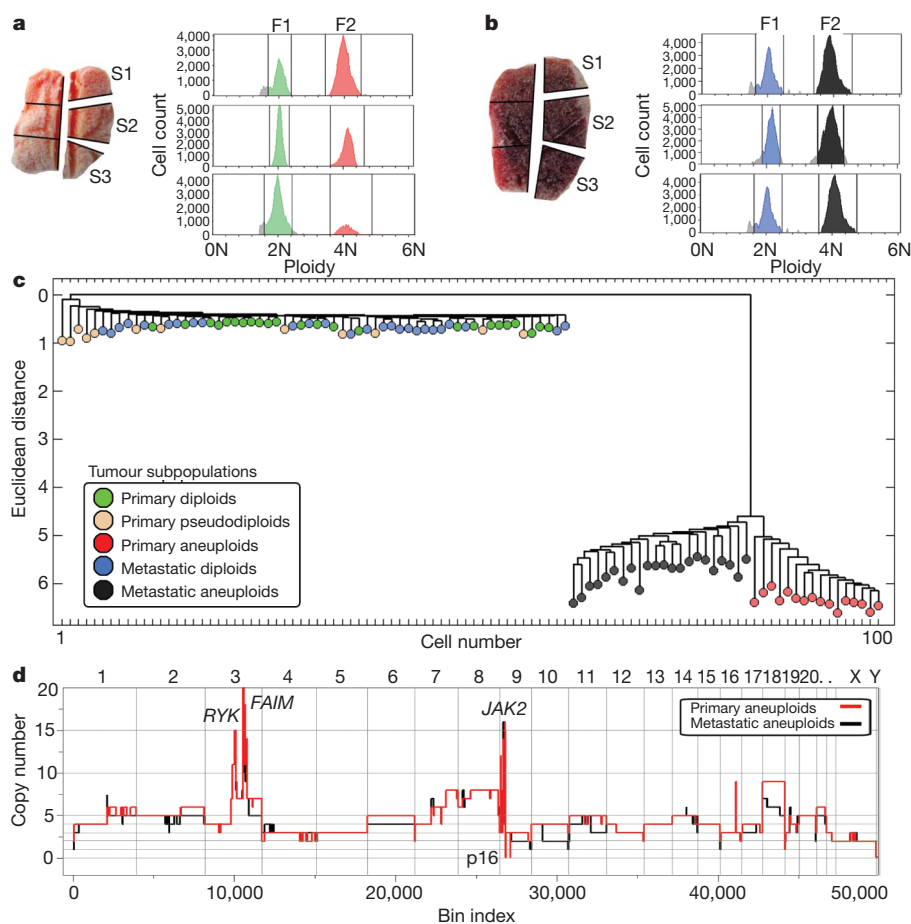


Figure 3 | Analysis of 100 single cells from a monogenic breast tumour and its liver metastasis. a, b, Primary breast tumour T16P was macrodissected and 52 nuclei were isolated from three sectors for FACS, showing two distributions of ploidy (F1 and F2). **b,** Liver metastasis T16M was macrodissected and 48 nuclei were isolated from three sectors for FACS also

showing two ploidy distributions (F1 and F2). **c,** Neighbour-joining tree of combined integer copy number profiles from the primary and metastatic tumours. **d,** Comparison of primary and metastatic aneuploid consensus copy number profiles.

major tumour subpopulations (Fig. 2c). This lineage shows that the n_1 common ancestor diverged a significant distance from the diploid cells, but that the distance between n_1 and n_2 is very small. By contrast, the divergence of the subpopulations after n_1 and n_2 is very large, with AB showing the greatest phylogenetic distance from the diploids. Thus we infer that the three subpopulations emerged when the tumour was much smaller.

We investigated a second tumour to determine whether these findings extend. We isolated 52 cells from a primary breast tumour (T16P) and 48 cells from its associated liver metastasis (T16M). Each tumour was macrodissected into six sectors, three of which were flow-sorted (Fig. 3a, b). Both T16M and T16P showed diploid peaks (F1) and a single aneuploid tetraploid peak (F2) of roughly equal cell count in all sectors (Supplementary Table 2), consistent with histological sections showing approximately 50% tumour and 50% normal (stromal) cells with low leukocyte infiltration in both samples. To explore population substructure we again constructed neighbour-joining trees from the integer copy number profiles, combining the primary and metastasis cells (Fig. 3c). We observed again numerous pseudodiploid cells, but a single subpopulation of aneuploid cells very diverged from the diploid population. As for T10, the 12 pseudodiploid cells from T16P showed diverse genomic lesions with no clear relationships to each other or to the main tumour lineage. Of the 24 normal diploids in the primary, two had deletions of the T-cell receptor. There were no pseudodiploid cells among the 26 diploid cells from the metastasis.

These data indicate that the primary tumour mass formed by a single clonal expansion of an aneuploid cell, and that one of the cells from this expansion subsequently seeded the metastatic tumour with little further evolution. There are no branches of the tree corresponding to cells intermediate between the aneuploid subpopulation and the diploid root. Although closely related, the primary and metastatic

aneuploid cells cleanly separate using the Euclidean metric (Fig. 3c), indicating that the two populations have not mixed since seeding the metastasis. The differences in the profiles that distinguish the primary and metastatic tumour populations are in the degree of copy number change rather than breakpoints (Fig. 3d). In a hierarchical tree created from breakpoints alone, we cannot cleanly separate primary from metastatic aneuploid cells (Supplementary Fig. 6b). Moreover, when we calculate common breakpoints in the single-cell profiles and apply biclustering to ordered samples (Supplementary Fig. 7b), a large number of breakpoints are common to both populations and no breakpoints cleanly distinguish them. By these analyses, no further population substructure is evident.

In contrast to the clear clonal relationships among aneuploid subpopulations, pseudodiploid cells are unusual in showing remarkable genomic heterogeneity (Fig. 4). Pseudodiploid profiles are characterized by nonrecurring copy number changes (including whole chromosome arms) that are not shared between any two pseudodiploid cells, nor with the corresponding tumour profiles (Fig. 4e). These data indicate that unlike the aneuploid cells, pseudodiploids do not undergo clonal expansions in the tumour. Nevertheless, they comprise a substantial proportion of the diploid gated cells: 8% in T10 (4/47) and 33% in T16P (12/36), or approximately 4% and 24% of the tumour mass, respectively. In contrast, the 18 profiles from single nuclei of normal adjacent breast tissue are all flat (Fig. 4a). The relative abundance of pseudodiploid cells in primary tumours indicates that they may emerge from an ongoing aberrant process that generates genomic diversity in the tumour.

In principle, we can learn about DNA sequence mutations from SNS data. However, the sparse sequence coverage makes this analysis problematic. By combining data from multiple cells, belonging to well-defined subpopulations, we can perform global and regional analysis at the many nucleotide positions where sufficient numbers of sequence

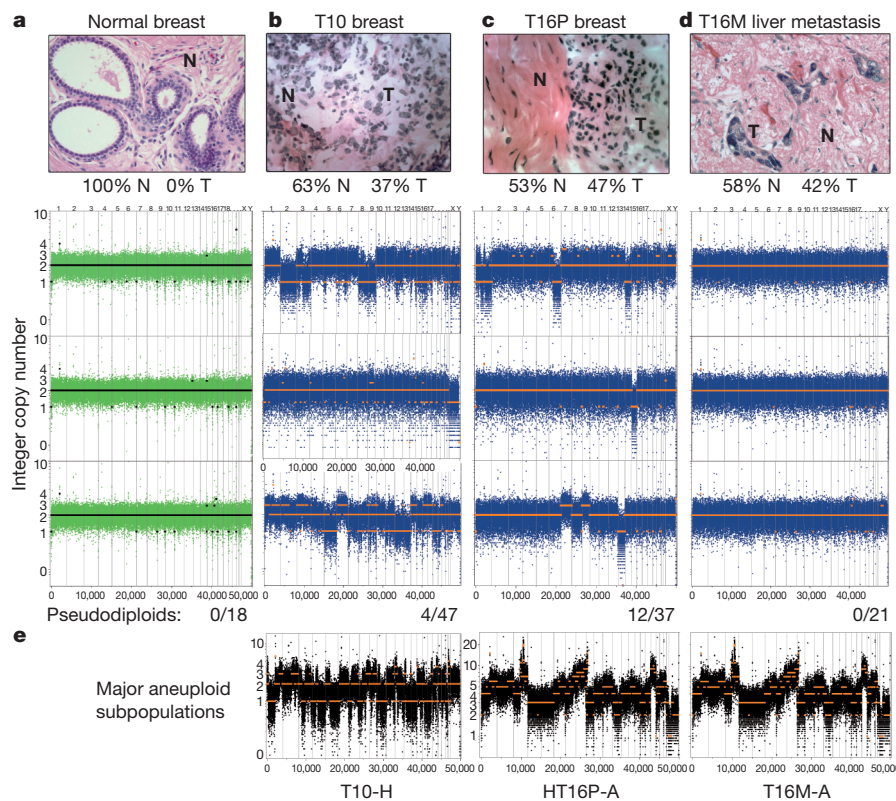


Figure 4 | Genetically diverse pseudodiploid cells in the diploid fractions of tumours. a–d, Haematoxylin and eosin stained tissues sections are shown in the upper panels with normal (N) and tumour (T) cell percentages indicated. Lower rows show bin counts and copy number profiles of single cells isolated from the 2N gated ploidy distributions, and the total number of cells analysed is

indicated below each column. The columns are: normal breast tissue cells (a); pseudodiploid cells in T10 (b); pseudodiploid cells in T16P (c); and diploid-gated nuclei from T16M (d). e, Bin counts and copy number profiles of single cells from the major aneuploid tumour subpopulations.

reads overlap. When examined this way, losses of heterozygosity are unequivocally significant, and map in large contiguous genomic blocks that correlate well with copy number loss (Supplementary Fig. 8 and Supplementary Table 3). The extensive loss of heterozygosity detected in all of the T10 subpopulations and in T16 indicates that both cancers passed through a hypodiploid stage.

Our study demonstrates that we can obtain robust high-resolution copy number profiles by sequencing a single cell and that by examining multiple cells from the same cancer we can make inferences about the evolution and spread of cancer. Moreover, the identification of pseudodiploid cells shows that these methods can identify cell types previously undetectable by other methods. Our findings are consistent with previous findings¹⁷ using bulk DNA, which indicate that copy number profiles in primary tumours are highly similar to the metastases. Thus, the metastatic cells emerge from a main advanced expansion, and not from an earlier intermediate or a completely different subpopulation. This is consistent with recent deep-sequencing studies of primary–metastatic pairs, all indicating that metastatic cells arise late in tumour development^{18,19}.

There are many gradual models for tumour progression, including clonal evolution²⁰, the mutator phenotype^{21,22} and stochastic progression²³. Although we have examined only two cancers in depth, both show a pattern of tumour growth that we call ‘punctuated clonal evolution’, borrowing a term from species evolution used to explain gaps in the fossil record²⁴. Explicitly, the tumour subpopulations are each distant from their root, without observable intermediate branching. In contrast to gradual models, this pattern reflects the sudden emergence of a tumour cell whose rate of effective population growth markedly exceeds its rate of genomic evolution.

METHODS SUMMARY

To perform SNS, nuclei are isolated either from cells in culture or frozen tumour sections and stained with 4',6-diamidino-2-phenylindole (DAPI). We use FACS to gate a desired population of nuclei by total DNA content and to deposit nuclei singly into 96-well plates. After WGA using Sigma GenomePlex, we sonicate to create free DNA ends without WGA adapters, and then construct libraries for 76 bp, single-end sequencing using one lane of an Illumina GA2 flowcell per nucleus. For each nucleus we typically achieve 9 million (mean = 9.042 million, s.e.m. \pm 0.328, n = 200) uniquely mapping reads using the Bowtie²⁵ alignment software. These sequences cover about 6% (mean = 5.95%, s.e.m. \pm 0.229, n = 200) of the genome, and are used to count sequence reads in 50,000 variable bins. The bin counts are segmented using a KS statistic and used to calculate integer copy number profiles. Neighbour-joining trees are constructed from the integer profiles and from the chromosome breakpoint patterns of each cell to infer evolution.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 May 2010; accepted 7 January 2011.

Published online 13 March 2011.

1. Park, S. Y., Gonen, M., Kim, H. J., Michor, F. & Polyak, K. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *J. Clin. Invest.* **120**, 636–644 (2010).
2. Torres, L. *et al.* Intratumor genomic heterogeneity in breast cancer with clonal divergence between primary carcinomas and lymph node metastases. *Breast Cancer Res. Treat.* **102**, 143–155 (2007).
3. Farabegoli, F. *et al.* Clone heterogeneity in diploid and aneuploid breast carcinomas as detected by FISH. *Cytometry* **46**, 50–56 (2001).

4. Chiang, D. Y. *et al.* High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nature Methods* **6**, 99–103 (2009).
5. Yoon, S., Xuan, Z., Makarov, V., Ye, K. & Sebat, J. Sensitive and accurate detection of copy number variants using read depth of coverage. *Genome Res.* **19**, 1586–1592 (2009).
6. Alkan, C. *et al.* Personalized copy number and segmental duplication maps using next-generation sequencing. *Nature Genet.* **41**, 1061–1067 (2009).
7. Geigl, J. B. *et al.* Identification of small gains and losses in single cells after whole genome amplification on tiling oligo arrays. *Nucleic Acids Res.* **37**, e105 (2009).
8. Fuhrmann, C. *et al.* High-resolution array comparative genomic hybridization of single micrometastatic tumor cells. *Nucleic Acids Res.* **36**, e39 (2008).
9. Pugh, T. J. *et al.* Impact of whole genome amplification on analysis of copy number variants. *Nucleic Acids Res.* **36**, e80 (2008).
10. Talseth-Palmer, B. A., Bowden, N. A., Hill, A., Meldrum, C. & Scott, R. J. Whole genome amplification and its impact on CGH array profiles. *BMC Res. Notes* **1**, 56 (2008).
11. Hughes, S. *et al.* Use of whole genome amplification and comparative genomic hybridization to detect chromosomal copy number alterations in cell line material and tumour tissue. *Cytogenet. Genome Res.* **105**, 18–24 (2004).
12. Huang, J., Pang, J., Watanabe, T., Ng, H. K. & Ohgaki, H. Whole genome amplification for array comparative genomic hybridization using DNA extracted from formalin-fixed, paraffin-embedded histological sections. *J. Mol. Diagn.* **11**, 109–116 (2009).
13. Navin, N. *et al.* Inferring tumor progression from genomic heterogeneity. *Genome Res.* **20**, 68–80 (2010).
14. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
15. Hicks, J. *et al.* Novel patterns of genome rearrangement and their association with survival in breast cancer. *Genome Res.* **16**, 1465–1479 (2006).
16. Prelic, A. *et al.* A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics* **22**, 1122–1129 (2006).
17. Liu, W. *et al.* Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nature Med.* **15**, 559–565 (2009).
18. Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999–1005 (2010).
19. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467**, 1114–1117 (2010).
20. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
21. Loeb, L. A., Springgate, C. F. & Battula, N. Errors in DNA replication as a basis of malignant changes. *Cancer Res.* **34**, 2311–2321 (1974).
22. Bielas, J. H., Loeb, K. R., Rubin, B. P., True, L. D. & Loeb, L. A. Human cancers express a mutator phenotype. *Proc. Natl Acad. Sci. USA* **103**, 18238–18242 (2006).
23. Heng, H. H. *et al.* Stochastic cancer progression driven by non-clonal chromosome aberrations. *J. Cell. Physiol.* **208**, 461–472 (2006).
24. Gould, S. J. & Eldredge, N. Punctuated equilibria comes of age. *Nature* **366**, 223–227 (1993).
25. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Ronemus, T. Spencer, A. Leotta, J. Meth, M. Kramer, L. Gelley, E. Ghiban. We also thank P. Blake and N. Navin at Sophic Systems Alliance. This work was supported by the NCI T32 Fellowship to N.N., and grants to M.W. and J.H. from the Department of the Army (W81XWH04-1-0477), the Breast Cancer Research Foundation, and the Simons Foundation. M.W. is an American Cancer Society Research Professor.

Author Contributions N.N. designed and performed experiments and analysis, and wrote the manuscript. J.K., A.K., L.M., D.L. and P.A. developed analysis programs. J.T., L.R., K.C., J.M., D.E. and A.S. performed experiments. W.R.M. designed experiments. J.H. and M.W. designed experiments, performed analysis and wrote manuscript.

Author Information All data has been deposited into the NCBI Sequence Read Archive under accession number SRA018951.105. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.W. (wigler@cshl.edu).

METHODS

Samples. The frozen ductal carcinoma T10 (CHTN0173) was obtained from the Cooperative Human Tissue Network, and T16P and T16M were obtained from Asterand. Pathology shows that both tumours were poorly differentiated and high grade (III) as determined by the Bloom–Richardson score, and triple-negative (ER[−], PR[−] and HER2/NEU[−]) as determined by immunohistochemistry. The cell lines used in this study include a normal male immortalized skin fibroblast (SKN1) and a breast cancer cell line (SK-BR-3). Normal breast tissue was obtained from H. Hibshoosh from Columbia University.

SNS. Nuclei were isolated from cell lines and from the frozen tumour using an NST-DAPI buffer (800 ml of NST (146 mM NaCl, 10 mM Tris base at pH 7.8, 1 mM CaCl₂, 21 mM MgCl₂, 0.05% BSA, 0.2% Nonidet P-40)), 200 ml of 106 mM MgCl₂, 10 mg of DAPI, and 0.1% DNase-free RNase A. The frozen tumour was first macrodissected into 12 sectors of equal size using surgical scalpels and nuclei were isolated from six sectors for FACS by finely mincing a tumour sector in a Petri dish in 1.0–2.0 ml of NST-DAPI buffer using two no. 11 scalpels in a cross-hatching motion. The cell lines were lysed directly in a culture plate using the NST-DAPI buffer, after first removing the cell culture media. All nuclei suspensions were filtered through 37-µm plastic mesh before flow-sorting.

Single nuclei were sorted by FACS using the BD Biosystems Aria II flow cytometer by gating cellular distributions with differences in their total genomic DNA content (or ploidy) according to DAPI intensity. First, a small amount of prepared nuclei from each tumour sample was mixed with a diploid control sample (derived from a lymphoblastoid cell line of a normal person) to accurately determine the diploid peak position within the tumour and establish FACS collection gates. Before sorting single nuclei, a few thousand cells were sorted to determine the DNA content distributions for gating. A 96-well plate was prepared with 10 µl of lysis solution in each well from the Sigma-Aldrich GenomePlex WGA4 kit. Single nuclei were deposited into individual wells in the 96-well plate along with several negative controls in which no nuclei were deposited.

WGA was performed on single flow-sorted nuclei as described in the Sigma-Aldrich GenomePlex WGA4 kit (catalogue no. WGA4-50RXN) protocol. WGA fragments from the frozen breast tumour and SK-BR-3 single cells were used directly for single-read library construction using the Illumina Genomic DNA Sample Prep Kit (catalogue no. FC-102-1001) and following standard protocol with a gel purification size range of 300–250 bp. WGA fragments from the fibroblast cell line were first sonicated using the Diagenode Bioruptor using the following program: 2 times, 7 min with 30 s high on/off mode in ice-cold water. Sonication removes a specific 28 bp adaptor sequence that is added on during WGA, and improves the total number of sequencing reads per lane.

Single-read libraries from single nuclei were sequenced on individual flow-cell lanes using the Illumina GA2 analyser for 76 cycles. Data was processed using the Illumina GAPIipeline-1.3.2 to 1.6.0. Sequence reads were aligned to the human genome (HG18/NCBI36) using the Bowtie alignment software²⁵ with the following parameters: ‘bowtie -S -t -m 1 -best -strata -p16’ to report only top scoring unique mappings for each sequence read. For each nucleus we typically achieve 9 million (mean = 9.042 million, s.e.m. ± 0.328, *n* = 200) uniquely mapping reads. These sequences cover about 6% (mean = 5.95%, s.e.m. ± 0.229, *n* = 200) of the genome uniquely. To eliminate PCR duplicates, we removed sequences with identical start coordinates.

Read depth counting in variable bins. Copy number is calculated from read density, by dividing the genome into ‘bins’ and counting the number of unique reads in each bin. In previous copy number studies read density was calculated using bins with uniform fixed length^{16–19}. In contrast, we use bins of variable length that adjust size depending on the mappability of sequences to regions of the human genome. In regions of repetitive elements, lower numbers of reads are expected and thus the bin size is increased. To determine interval sizes we simulated sequence reads by sampling 200 million sequences of length 48 from the human reference genome (HG18/NCBI36) and introduced single nucleotide errors with a frequency encountered during Illumina sequencing. These sequences were mapped back to the human reference genome using Bowtie²⁵ with unique parameters as described earlier. We assigned a number of bins to each chromosome based on the proportion of simulated reads mapped. We then divided each chromosome into bins with an equal number of simulated reads. This resulted in 50,009 genomic bins with no bins crossing chromosome boundaries. The median genomic length spanned by each bin is 54 kb. For each cell the number of reads mapped to each variable length bin was counted. This variable binning efficiently reduces false deletion events when compared to uniform length-fixed bins as shown in Supplementary Fig. 2b and c. For a single cell we typically measure 138 sequence reads per bin.

Integer copy number quantification. Single cells will have integer copy number states that we can infer from sequence read counts, as follows. Unique sequence reads are counted in variable bins (Supplementary Fig. 4a) and segmented using

the Kolmogorov–Smirnov (KS) statistic (Supplementary Fig. 4b). To estimate the integer differences of copy number states, we calculate Gaussian kernel smoothed density plots using Splus (MathSoft), showing the difference between median bin counts for all pair-wise combinations of different segments (Supplementary Fig. 4c–e). The uniform steps between groups are very apparent, and are a general property of single-cell data. We then convert our KS-segmented data into profiles of integer copy number as follows. We take the differential bin count of the second peak, denoted by an asterisk in Supplementary Fig. 4a, to represent a copy number ‘increment’ of 1. We then divide every bin count in the profile by the increment and round to infer the integer copy number. We show in Supplementary Fig. 4f–g how closely the segmentation profile agrees with the integer copy number profile. However, for diploid or near diploid cells there are few to no steps from which to observe the increment, and we use a different method, taking the increment as the median bin count on the autosomes divided by two.

Gene annotations. Amplifications and deletions identified in the single-cell copy number profiles were annotated to identify UCSC genes. Cancer genes were identified using a compiled database from the cancer gene consensus and the NCI cancer gene index (Sopich Systems Alliance, Biomax Informatics AG).

Neighbour-joining trees of copy number profiles. Integer copy number profiles of single cells were used to calculate neighbour-joining trees using a Euclidean distance metric with Matlab (Mathworks). Branches were flipped to orient nodes within subpopulations and trees were rooted using the last common diploid node.

Common breakpoint detection. Breakpoints are defined as bins with a copy number different than the previous bin in genome order. A transition from a lower copy number to a higher copy number (in genome order) is considered to be a different event than the opposite transition. To find breakpoint regions we count each breakpoint in each cell and the immediately neighbouring bins. A contiguous set of bins with counts greater than 1 is designated a breakpoint region. This results in a set of common breakpoint regions. Each cell is then scored for the occurrence of each of these events, a one meaning the cell has a copy number transition of that type (low to high or high to low) in that genomic region and a zero meaning no copy number transition of that type in that region.

Hierarchical tree of chromosome breakpoints. We used chromosome breakpoints patterns to build a neighbour-joining tree. To eliminate breakpoint events with a high standard deviation, we limited our analysis to breakpoint regions covering no more than seven adjacent bins (*N* = 657). Using a Euclidean metric, we calculated a distance matrix from the binary chromosome breakpoint patterns identified in the single cells using Matlab (Mathworks). From this distance matrix we constructed a tree using average linkage.

Heatmap of chromosome breakpoints. The biclustering heatmap is based on the same set of breakpoints used to build the neighbour-joining tree. Colour indicates the presence of an event, and white means no event. The columns are ordered as in the tree. The rows are events ordered to show clearly which of the subsets of the four main groups share which events. The groups are ordered by subpopulation. A four-dimensional binary vector represents each of the 16 possible subsets of these groups (subset vector). Each breakpoint is represented by a four-dimensional vector of the per cent of cells in each group having an event at that breakpoint (the ‘breakpoint vector’). The angle from each breakpoint vector to each subset vector is computed as well as the length of each projection vector. If the length of the projection vector is less than 0.05 the breakpoint vector is assigned to the empty (0,0,0,0) subset, otherwise it is assigned to the subset vector with the smallest angle to the breakpoint vector. The rows are ordered by subset vector in the following order: (1,1,1,1), (0,0,0,1), (0,1,0,0), (1,0,0,0), (0,0,1,1), (0,1,0,1), (1,0,0,1), (0,1,1,0), (1,0,1,0), (1,1,0,0), (0,1,1,1), (1,0,1,1), (1,1,0,1), (1,1,1,0), (0,0,0,0). Within each subset the rows are in descending order by the number of cells in that subset having that event and then in ascending order by the number of cells outside of that subset that do not have that same event.

Analysis of loss of heterozygosity using sequence mutations. PCR duplicates were removed from mapped sequence reads and bases with a quality score below 30 were excluded from analysis. We then determined the set of observed nucleotide types for each cell sequenced from the T10 and T16P and T16M tumours and every position in the genome. For each subpopulation we classified a position as the observed nucleotides only if one or two nucleotide types were each observed in five or more cells in the subpopulation. For each grouping of subpopulations DH, DA, if a classification was made in every subpopulation in the group, we translated the classifications into the generic nucleotides (a,b) based upon the order in which they were seen in the group, from left to right. We counted the resulting classifications of positions for each group by class, and determined whether long blocks of identical classifications along a chromosome were expected by chance. To establish the significance of our classification counts, we repeated our analysis 100 times with randomly permuted cell labels within each group of subpopulations. We eliminated any effects from differing subpopulation size in a separate set of runs of the same analysis, each with 24 randomly selected cells in every subpopulation.

Catecholamine receptor polymorphisms affect decision-making in *C. elegans*

Andres Bendesky¹, Makoto Tsunozaki¹, Matthew V. Rockman², Leonid Kruglyak³ & Cornelia I. Bargmann¹

Innate behaviours are flexible: they change rapidly in response to transient environmental conditions, and are modified slowly by changes in the genome. A classical flexible behaviour is the exploration–exploitation decision, which describes the time at which foraging animals choose to abandon a depleting food supply. We have used quantitative genetic analysis to examine the decision to leave a food patch in *Caenorhabditis elegans*. Here we show that patch-leaving is a multigenic trait regulated in part by naturally occurring non-coding polymorphisms in *tyra-3* (tyramine receptor 3), which encodes a G-protein-coupled catecholamine receptor related to vertebrate adrenergic receptors. *tyra-3* acts in sensory neurons that detect environmental cues, suggesting that the internal catecholamines detected by *tyra-3* regulate responses to external conditions. These results indicate that genetic variation and environmental cues converge on common circuits to regulate behaviour, and suggest that catecholamines have an ancient role in regulating behavioural decisions.

Despite abundant evidence for heritability of behavioural traits within and between species, only a few naturally varying traits have been associated with polymorphisms in specific genes¹. Foraging for food is an ecologically relevant, environmentally regulated behaviour that is suitable for genetic analysis, as it can differ between populations of a species that live in different habitats². An essential foraging decision is the choice between exploiting existing resources and exploring other options that may provide new resources. This decision can be described by Charnov's marginal value theorem, which proposes that the optimal time for an animal to leave a foraging ground occurs when local resource levels fall below the average level in the entire habitat³. The marginal value theorem was developed for animals foraging for food in patchy environments, but has analogies with diverse decision-making processes in field biology, cognitive neuroscience and economics^{2,4–6}.

Studies of patch-leaving behaviour in the nematode *C. elegans* have revealed innate, environmental and experience-dependent factors that affect its foraging decisions. *C. elegans* rarely leaves a dense lawn of high-quality bacterial food^{7,8}, but more frequently leaves lawns of pathogenic bacteria or lawns that are spiked with chemical repellents^{9,10}. Males will leave lawns that do not contain potential mates¹¹, while hermaphrodites leave lawns when animal density is high¹². In addition, wild-type strains vary in their propensity to leave bacterial lawns based on a genetic polymorphism that affects the G-protein-coupled neuropeptide receptor NPR-1 (refs 12–14). This *npr-1* polymorphism affects many foraging behaviours; low-activity *npr-1* strains aggregate into social feeding groups, move quickly on food, and have altered responses to oxygen, carbon dioxide and pheromones compared to the N2 laboratory strain^{15–20}. The high-activity allele of *npr-1* in N2 arose in the laboratory, probably as an adaptation to laboratory conditions¹⁹, so it is not known whether genetic variation affects *C. elegans* foraging in natural environments.

Natural genetic variation within a species can generate diversity in foraging behaviour, as exemplified by the polymorphic *Drosophila melanogaster* foraging (*for*) gene, which encodes a cyclic guanosine

monophosphate (cGMP)-dependent protein kinase²¹. A low-activity allele of *for* is present in *Drosophila* sitter larvae, which move slowly on a food patch; a high-activity allele of *for* is present in rover larvae, which move quickly and disperse rapidly²². A *for*-related cGMP-dependent kinase affects foraging in honeybees, ants and nematodes, suggesting that diverse animals share molecular mechanisms for behavioural regulation^{22,23}.

To gain further insight into the genetics and neurobiology of exploratory behaviour in *C. elegans*, we here use quantitative genetic analysis to examine this behaviour's genetic architecture in wild-type strains, and show that genetic variation in multiple loci, including a catecholamine receptor, interacts with environmental conditions to regulate the exploitation–exploration decision.

Multiple loci affect leaving behaviour

Different wild-type strains of *C. elegans* vary in their tendency to leave or remain on a standardized small lawn of bacterial food (Fig. 1a). For example, adult hermaphrodites from the laboratory strain N2 leave the lawn only once every 100 min, whereas animals from the CB4856 (HW) strain isolated from pineapple fields in Hawaii leave the lawn once every 5–6 min (Fig. 1b, Supplementary Movies 1 and 2). To determine the genetic architecture of this behavioural difference between N2 and HW, we quantified leaving rates in 91 N2-HW recombinant inbred advanced intercross lines (RIAILs)²⁴. 58 of the RIAILs had low leaving rates comparable to N2, only 6–10 had high leaving rates comparable to HW, and 23 had intermediate rates (Fig. 1c). The excess of low leaving rates and the continuous behavioural distribution in RIAILs suggest that leaving is a multigenic quantitative trait.

Quantitative trait locus (QTL) analysis of the RIAILs uncovered two regions with significant effects on leaving rates, one on the X chromosome and one on chromosome II (Fig. 1d). The X chromosome QTL overlapped with the location of the polymorphic G-protein-coupled neuropeptide receptor NPR-1, which affects many food-related behaviours^{12,15}. The *npr-1* polymorphism has previously

¹Howard Hughes Medical Institute, Laboratory of Neural Circuits and Behavior, The Rockefeller University, New York, New York 10065, USA. ²Department of Biology and Center for Genomics and Systems Biology, New York University, New York, New York 10003, USA. ³Howard Hughes Medical Institute, Lewis-Sigler Institute for Integrative Genomics and Department of Ecology and Evolutionary Biology, Carl Icahn Laboratory, Princeton University, Princeton, New Jersey 08544, USA.

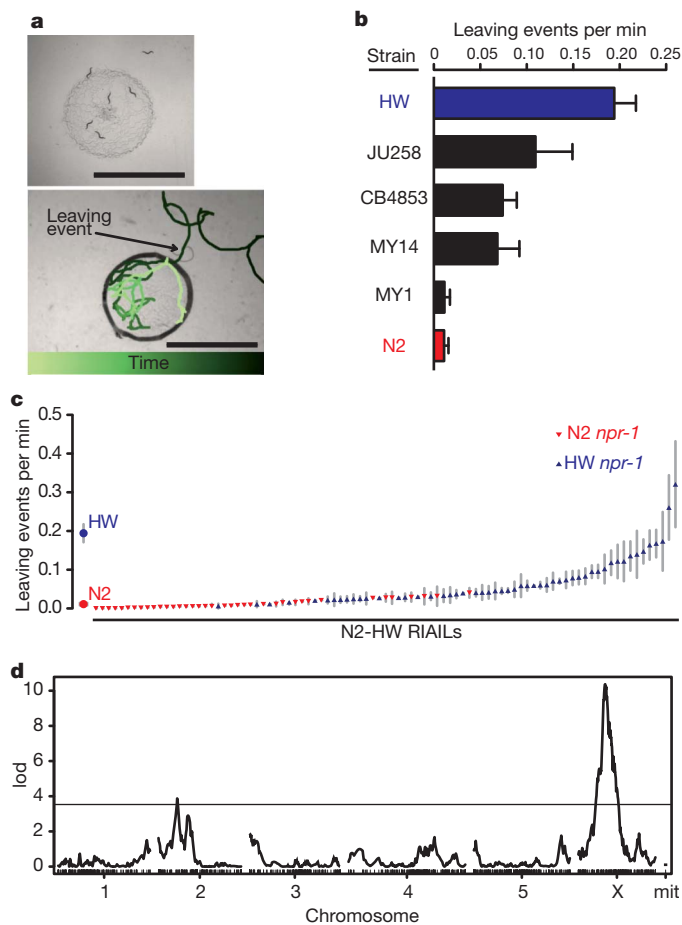


Figure 1 | Lawn-leaving behaviour varies between wild-type *C. elegans* strains. **a**, Lawn-leaving assays. Top, six adult HW hermaphrodites on a bacterial lawn. One animal has left the lawn and one is leaving. Bottom, track of a HW animal during 5 min of an assay. The track is colour-coded to show passage of time (colour scale at bottom). The border of the lawn is outlined. Scale bar, 6 mm. **b**, Leaving rates of six wild-type strains. **c**, Leaving rates of 91 N2-HW recombinant inbred advanced intercross lines (RIAILs)²⁴ and parental strains. **d**, QTL analysis of RIAILs shown in **c**. The horizontal line denotes the $P < 0.01$ genome-wide significance threshold. lod, log likelihood ratio. Error bars in **b** and **c**, s.e.m.

been shown to affect leaving, as well as locomotion speed on food, a behaviour that partially correlates with leaving rate^{12,15} (Supplementary Fig. 2). Examining the *npr-1* genotype in the RIAILs revealed a strong but asymmetric correlation with leaving rates (Fig. 1c). Every strain with the N2 allele of *npr-1* had low leaving rates (≤ 1 event every 20 min), but strains with the HW allele of *npr-1* could have either low or high leaving rates (Fig. 1c). The asymmetric distribution is consistent with a role for *npr-1* in leaving behaviour, but indicates that *npr-1* has epistatic interactions with other loci segregating in the RIAILs.

The involvement of *npr-1* in leaving behaviour was confirmed by analysing near-isogenic lines (NILs) containing the N2 and HW *npr-1* alleles in the reciprocal strain background, and by examining *npr-1* null mutants (Supplementary Fig. 3). Specific transgenic expression of the N2 *npr-1* allele in its essential site of action, the RMG motor neurons²⁰, sharply reduced the leaving rate of HW animals (Supplementary Fig. 3). Thus *npr-1* is a regulator of HW leaving rates, but not the only contributing gene.

tyra-3 affects leaving behaviour

Studies in yeast, flies, mice, and plants have shown that individual QTLs often resolve into several genes that contribute to phenotypic variance^{25–28}. Similarly, fine-mapping of the ~ 1 megabase (Mb) QTL

that contained *npr-1* suggested the existence of multiple loci that affected leaving rates. A NIL with <150 kilobases (kb) of N2 DNA spanning the *npr-1* locus introgressed into HW had N2-like leaving rates (*leav-1* QTL, Fig. 2a and Supplementary Fig. 3). A second NIL with 700 kb of N2 DNA that did not cover *npr-1* introgressed into HW also had a low leaving rate, with about half the leaving rate of HW (*leav-2* QTL, Fig. 2a). These results suggest the existence of a second X-linked locus that affects leaving rates, which we called *leav-2*. The *leav-2* region did not affect leaving in the N2 genetic background (Fig. 2a), so all subsequent experiments were conducted in the HW background.

A 100 kb minimal region for *leav-2* was identified by analysing the breakpoints of individual RIAILs (Supplementary Fig. 4 and Supplementary Methods). We characterized the genetic properties of *leav-2* by crossing the *leav-2* NIL strain with HW. The heterozygous F₁ progeny had leaving rates similar to the *leav-2* NIL (Fig. 2a), indicating that the N2 *leav-2* locus was dominant to HW and suggesting that N2 transgenes covering the relevant gene should reduce the leaving rate of HW animals. Therefore, overlapping N2 genomic DNA fragments from the 100 kb minimal *leav-2* region were introduced into HW animals by microinjection (Fig. 2b and Supplementary Fig. 5). A single gene in this region reduced leaving rates: *tyra-3*, which encodes a G-protein-coupled receptor for the invertebrate noradrenaline-like neurotransmitters tyramine and octopamine²⁹. Tyramine and octopamine receptors are related to vertebrate adrenergic receptors, and are thought to carry out analogous functions. *tyra-3* genomic fragments from the N2 strain were more active than *tyra-3* fragments from the HW strain injected at the same concentration, consistent with the possibility that *tyra-3* is a polymorphic gene that differs between N2 and HW (Fig. 2b).

If *leav-2* corresponds to *tyra-3*, a *tyra-3* mutation should eliminate its activity³⁰. To test this prediction genetically, a null allele of *tyra-3* in an N2 background was introgressed into a HW background. The N2 region in the resulting NIL covered from 4.9 to 5.4 Mb of the X chromosome, the inferred position of *leav-2*. The *tyra-3(ok325)* null NIL had high (HW-like) leaving rates, suggesting that N2 *leav-2* activity was not present in the strain (Fig. 2a). Heterozygotes between HW and the near-isogenic *tyra-3(ok325)* null strain also had high leaving rates (Fig. 2a). These results are as expected if the active locus in *leav-2* is *tyra-3*; however, other genes within the introgressed regions could also contribute to the different leaving rates.

To strengthen the connection between *tyra-3* and *leav-2*, RNA interference against *tyra-3* was performed in the *leav-2* NIL that has low leaving rates due to the presence of the N2 QTL. Knockdown of *tyra-3* increased the leaving rate of the *leav-2* NIL to levels observed in HW animals, the result predicted if the *tyra-3* locus from N2 reduces leaving (Fig. 2c). Comparable experiments in a pure HW strain had minimal effects, as expected if *tyra-3* activity in HW is already low.

Further confirmation that the HW allele of *tyra-3* has reduced biological activity was provided by examining the one phenotype previously associated with *tyra-3*, avoidance of dilute octanol²⁹. *tyra-3* null mutants avoid octanol more strongly than wild-type N2; the NIL strain with the HW *tyra-3* allele had a similar enhanced octanol response, suggesting that the HW *tyra-3* allele has reduced *tyra-3* function (Supplementary Fig. 6).

Non-coding changes affect *tyra-3* activity

The differential activity of N2 and HW genomic *tyra-3* fragments in the leaving assay suggested that N2 and HW alleles are functionally distinct (Fig. 2b). To identify polymorphisms between N2 and HW alleles of *tyra-3*, we sequenced ~ 19 kb surrounding the *tyra-3* locus in HW. There were 34 differences between HW and the N2 consensus genomic sequence (Fig. 3a): 33 non-coding changes and a single coding difference that changed a glutamic acid in the *tyra-3b* isoform to glycine.

Sequences that contribute to the differential activity of N2 and HW *tyra-3* alleles were localized further using transgenic assays. We fused

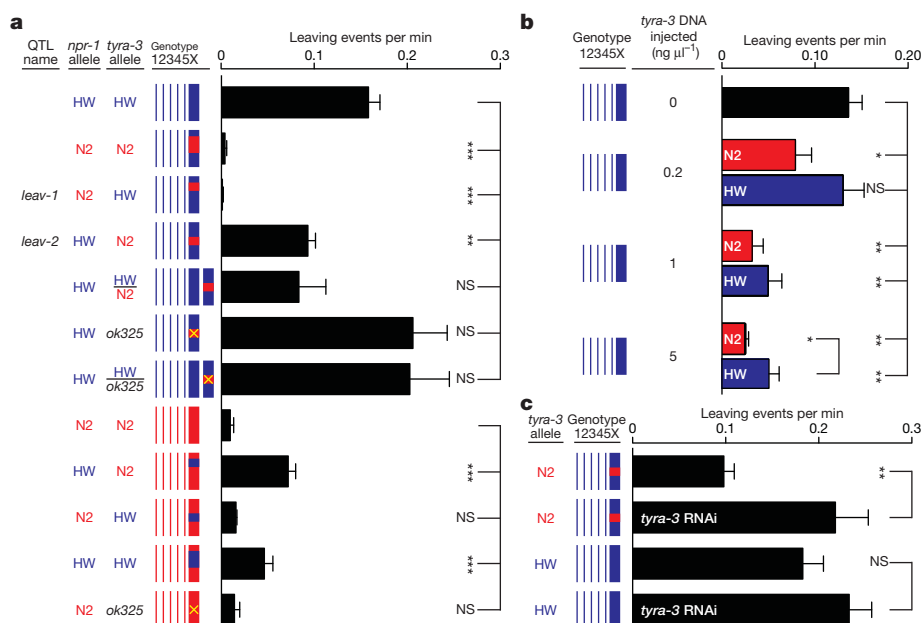


Figure 2 | N2 and HW *tyra-3* alleles differentially affect leaving rates.

a, Dissection of the QTL on X into two loci: *leav-1* (4.70–4.78 Mb) and *leav-2* (4.78–5.75 Mb). ‘Genotype’ shows chromosomes; thick line is X chromosome. Blue denotes HW DNA, red denotes N2 DNA, and yellow denotes the *tyra-3(ok325)* null mutant. In heterozygous strains, both X chromosomes are shown.

N2 and HW *tyra-3b* complementary DNAs to 4.9 kb of non-coding N2 or HW sequence upstream of the *tyra-3b* start site and introduced each of the four resulting clones into the HW strain. *tyra-3* transgenes with the N2 non-coding sequence were significantly more potent than comparable transgenes with the HW sequence, regardless of whether they preceded N2 or HW *tyra-3* cDNAs (Fig. 3b), excluding the coding polymorphism and localizing a functional difference between N2 and HW *tyra-3* genes to a 4.9 kb region that harbours 5 non-coding SNPs, 1 single nucleotide insertion, and a 184 bp deletion in HW. To

b, *tyra-3* genomic fragments (Fig. 3a) reduce HW leaving rates. Blue, HW transgenes; red, N2 transgenes. Two-way analysis of variance (ANOVA) showed significant effects of both transgene concentration and DNA strain of origin. **c**, Effect of *tyra-3* RNAi. Error bars, s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ by *t*-test or ANOVA with Dunnett test. NS, not significant.

narrow the relevant change down further, the 184 bp deletion was engineered into the N2 *tyra-3* genomic fragment; this clone was significantly less potent in the leaving assay than the full N2 genomic fragment (Supplementary Fig. 7). These results indicate that the 184 bp deletion represents at least part of the functional difference between N2 and HW *tyra-3* alleles.

Sequence variation in *tyra-3* non-coding regions could affect the level or location of *tyra-3* expression. Quantitative reverse transcription-PCR analysis (RT-PCR) of *tyra-3* mRNA levels in mixed-stage animals

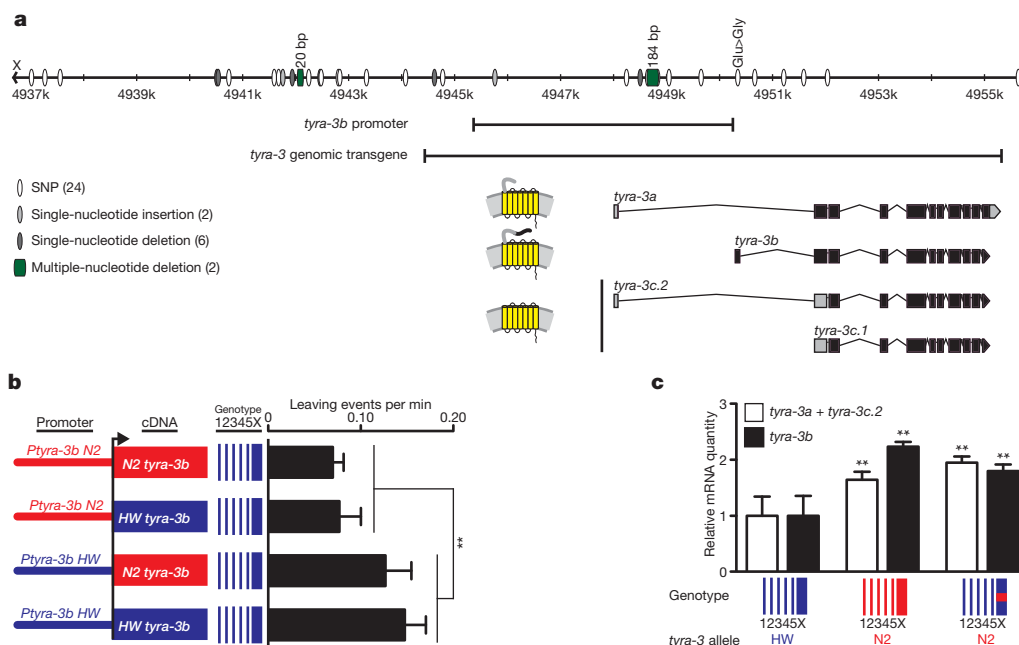


Figure 3 | Non-coding changes in *tyra-3* affect its activity and expression level. **a**, HW polymorphisms in the *tyra-3* locus relative to N2. *tyra-3* encodes three predicted G-protein-coupled receptors, shown as diagrams to the left of the gene structures for *tyra-3a*, *tyra-3b* and *tyra-3c*. The genomic region examined in Fig. 2b and the 4.9 kb promoter used in **b** and Fig. 4a are indicated.

b, Leaving rates of transgenic HW animals with *tyra-3b* promoters (*Ptyra-3b*) fused to *tyra-3b* cDNAs. Error bars, s.e.m. ****** $P < 0.01$ by two-way ANOVA; no statistical interaction between the promoter and the cDNA. **c**, Relative amounts of *tyra-3* isoform mRNAs (**a**) in HW, N2 and *leav-2* strains (Fig. 2a). Error bars, s.d. ****** $P < 0.01$ compared to HW, ANOVA with Dunnett test.

indicated that N2 expressed approximately twice as much *tyra-3* mRNA as HW, consistent with increased *tyra-3* activity in the N2 strain (Fig. 3c). The *leav-2* NIL with N2 *tyra-3* introgressed into HW also had high *tyra-3* messenger RNA levels, suggesting that *cis*-acting changes affect *tyra-3* expression (Fig. 3c).

As both N2 and HW were cultivated in the laboratory for many years before permanent cultures were frozen, we wished to exclude the possibility that the *tyra-3* polymorphisms were laboratory-derived¹⁹. Therefore, 19 kb of the *tyra-3* locus was sequenced in all wild strains tested for leaving behaviour in Fig. 1, including three strains that were frozen immediately after their isolation. Each strain represents a different *C. elegans* haplotype group²⁴. Both N2-like and HW-like *tyra-3* sequences were represented in the wild-caught strains, confirming the wild ancestry of both alleles (Supplementary Table 1 and Supplementary Methods). Notably, the *tyra-3* locus of MY1 was identical to N2 and, correspondingly, the leaving rate of MY1 was similar to that of N2.

tyra-3 acts in sensory neurons

The identification of *tyra-3* provided an opportunity to characterize the neuronal basis of the decision to leave or remain on a food patch. The biological activity of a transgene with 4.9 kb upstream of the *tyra-3b* start site fused to a *tyra-3* cDNA (Fig. 3b) implied that it was expressed in cells that regulate leaving behaviour. When this 4.9 kb region was fused to green fluorescent protein (GFP), it drove reliable expression in ASK, ADL, AIM, AUA, BAG, CEP, OLQ and SDQL neurons, in other unidentified neurons in the ventral ganglion and the tail, occasionally in AFD and AWC neurons, and in two non-neuronal cell types, the spermatheca and the distal tip cell (Fig. 4a and data not shown). The same set of cells was observed with reporter genes bearing either N2 or HW *tyra-3* upstream regions, and in both N2 and HW genetic backgrounds. Together with the quantitative RT-PCR data (Fig. 3c), these results suggest that different *tyra-3* expression levels, not different sites of expression, distinguish N2 and HW alleles.

The neurons whose activity is regulated by *tyra-3* were localized further by expressing *tyra-3* cDNAs from cell-type-specific promoters. *tyra-3* expression in ASK or BAG sensory neurons significantly reduced leaving, but expression in the CEP or ADL sensory neurons

did not (Fig. 4b). The ASK neurons sense attractive food-derived amino acids³¹ and regulate search behaviours after animals are removed from food^{32,33}. The BAG neurons sense CO₂ and O₂, two cues associated with bacterial metabolism^{34,35}. Lowering O₂ to levels that activate BAG reduced leaving rates (Supplementary Fig. 8).

To investigate whether the *tyra-3* non-coding polymorphism affects expression in relevant neurons, single-copy N2 or HW *tyra-3b* promoters driving GFP were inserted into a single, defined chromosomal location using the MosSCI technique³⁶. GFP levels in ASK neurons were significantly higher for transgenes containing the N2 promoter compared to those containing the HW promoter (Fig. 4c). These results suggest that the N2 *tyra-3* locus is associated with higher *tyra-3* expression in ASK, as well as higher *tyra-3* mRNA expression at a whole-animal level; expression in BAG was not examined.

The behavioural functions of ASK and BAG, and the effect of *tyra-3* on those functions, were assessed by killing the neurons in different genetic backgrounds. Killing the ASK neurons reduced the leaving rate of HW animals, indicating that ASK can promote leaving (Fig. 4d). The ablation resembled the effect of the ASK::*tyra-3* transgene, suggesting that *tyra-3* reduces ASK activity. In agreement with this idea, killing the ASK neurons in a strain with the N2 high-activity *tyra-3* allele did not reduce their leaving rates further. The effect of *tyra-3* on ASK was selective for this assay; *tyra-3* did not reduce lysine chemotaxis, a second ASK-dependent behaviour (Supplementary Fig. 9).

Killing the BAG neurons increased leaving rates in the strain with the N2 *tyra-3* allele, demonstrating that BAG neurons prevent leaving (Fig. 4d). However, killing BAG had no effect in the strain with the HW *tyra-3* allele, suggesting that BAG activity is already low in this strain under the assay conditions. The ablation and genetic results suggest that the N2 *tyra-3* allele decreases ASK activity and increases BAG activity, two changes that act together to prevent leaving (Supplementary Fig. 1).

Gene-gene-environment interactions

Like most natural behaviours, the decision to leave a food patch is regulated by multiple genes and the environment; it responds to genetic variation in *tyra-3*, *npr-1*, and additional genes on the autosomes

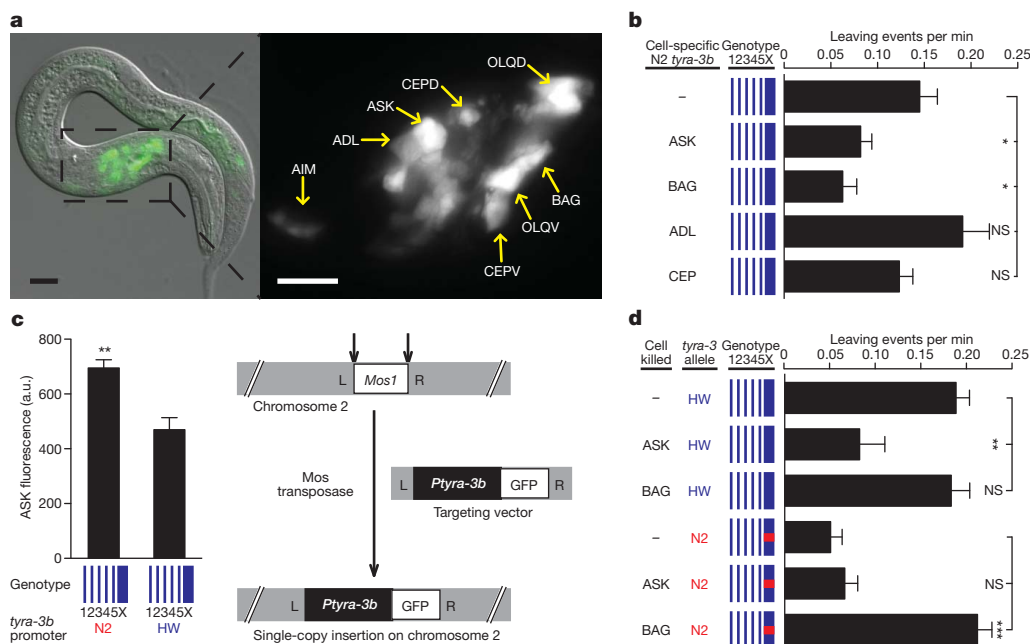


Figure 4 | *tyra-3* acts in ASK and BAG sensory neurons. **a**, Expression of 4.9 kb N2 *tyra-3b* promoter::GFP fusion (Fig. 3a) in HW animal at low (left) and high (right) magnification. HW *tyra-3b* promoter::GFP is expressed in the same cells, which are labelled in the right panel. Posterior signal is gut autofluorescence. Scale bar, 20 μ m. **b**, Leaving rates of HW strains expressing

tyra-3b in specific cells. **c**, Left, GFP fluorescence intensity in ASK of HW animals with a MosSCI insertion of N2 or HW 4.9 kb *tyra-3b* promoter::GFP. Right, schematic of MosSCI technique³⁶. **d**, Leaving rates after killing ASK or BAG in HW and *leav-2* strains (Fig. 2a). Error bars, s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ by *t*-test or ANOVA with Dunnett test.

(Fig. 1) as well as food quality and quantity^{7,8}. Our results suggested that the N2 *npr-1* allele was epistatic to *tyra-3*; animals with the N2 *npr-1* allele had low leaving rates regardless of the *tyra-3* genotype (Fig. 2a). However, N2 *npr-1* reduced the leaving rate to almost zero, making it difficult to detect any further reduction. To make the assay more powerful, leaving was assayed on bacterial lawns of different densities. Leaving rates of all genotypes increased on thinner lawns and decreased on thicker lawns (Supplementary Fig. 10), but the thickness of the lawn changed the genetic interaction between *tyra-3* and *npr-1*. In the standard leaving assay, *tyra-3* polymorphisms had different effects only in the presence of the HW *npr-1* allele; on a thinner lawn, only in the presence of the N2 *npr-1* allele (Supplementary Fig. 10). Thus the epistatic relationship between *npr-1* and *tyra-3* is defined by the specific environment, not by an intrinsic regulatory relationship between the genes.

Discussion

Our results show that natural variation in *tyra-3* affects patch leaving, a behaviour representative of the exploration–exploitation decision. *tyra-3* encodes a G-protein-coupled receptor activated by the invertebrate transmitters tyramine and octopamine²⁹, which are structurally related to vertebrate adrenaline and noradrenaline. Catecholamines are known to regulate arousal systems that affect many behaviours and behavioural decisions. In *C. elegans*, octopamine drives sensory, molecular and behavioural responses to starvation, and tyramine affects specific aspects of locomotion^{37–40}. In insects, octopamine affects locomotory activity, arousal and aggression^{41–43}. Mammalian noradrenaline is generally implicated in arousal behaviours, and noradrenaline release from the locus coeruleus is associated with switching between different tasks, a cognitive function with analogies to the exploration–exploitation decision⁶.

Relatively few natural behavioural variations have been mapped to the single-gene level in any animal, and it is interesting that several of these variations affect G-protein-coupled receptor signalling systems^{44,45}. We speculate that these receptor pathways may serve as common substrates of behavioural variation. All animal genomes encode many G-protein-coupled receptors with different expression patterns. These receptors may provide a reservoir for genetic changes, as alteration in an individual receptor could cause relatively discrete effects without disrupting the entire system.

QTL mapping in rodents and in *Drosophila* indicates that most behavioural traits are polygenic, with widespread epistatic effects^{1,46}. In agreement with this conclusion, our analysis suggests the existence of epistatic interactions between *tyra-3*, *npr-1* and at least one additional locus. Importantly, the non-additive interactions between *tyra-3* and *npr-1* are not stable, but vary based on the genetic background and the environment, similar to what has been found with yeast sporulation QTLs⁴⁷.

By integrating genetic studies of *C. elegans* foraging with neuronal analysis, we can provide a first-level description of underlying mechanisms. The sensory neurons that express *tyra-3* detect food-related cues; we suggest that they integrate these external cues with internal arousal states detected by *tyra-3*, and that different *tyra-3* alleles confer differential sensitivity to these arousal states (Supplementary Fig. 1). Thus variation in *tyra-3* lies at the intersection of many forms of behavioural flexibility: rapid responses to environmental cues, short-term modulation by internal state fluctuations, and long-term genetic changes that lead to adaptive changes in innate behaviours.

METHODS SUMMARY

Standardized leaving assays were conducted by videotaping seven adult hermaphrodite animals for 30 min on a bacterial lawn that was low-density compared to their growth lawn. The number of leaving events was recorded manually by examining the video recordings, and further behavioural analysis was conducted with a Matlab code adapted from the Parallel Worm Tracker⁴⁸. A leaving event was defined as an

episode in which the whole body of an animal left the bacterial lawn and the animal did not reverse immediately to return to the lawn. The leaving rate was calculated as the number of leaving events per worm minute spent inside the bacterial lawn. Experiments on each strain were repeated at least three times.

QTL analysis was performed on the mean leaving rates of N2–HW recombinant inbred advanced intercross lines (RIAILs) by non-parametric interval mapping in R/qtl^{24,49}. Significance levels were estimated from 10,000 permutations of the data. Near-isogenic lines (NILs) were created by backcrossing a chromosomal region or allele into the desired genetic background at least 9 times. Extrachromosomal transgenes were made by injection of DNA clones into the gonads of young adult hermaphrodites together with a fluorescent coinjection marker⁵⁰. To control for variation between transgenes, two to five independent lines from each injection were characterized. Single-copy insertion of transgenes was performed using the direct MosSCI transposition technique, targeting the *ttTi5605* Mos allele on chromosome II³⁶.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 March 2010; accepted 14 January 2011.

Published online 16 March 2011.

1. Flint, J. & Mackay, T. F. Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res.* **19**, 723–733 (2009).
2. Stephens, D. W., Brown, J. S. & Ydenberg, R. C. *Foraging: Behavior and Ecology* (Univ. Chicago Press, 2007).
3. Charnov, E. L. Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* **9**, 129–136 (1976).
4. March, J. G. Exploration and exploitation in organizational learning. *Organ. Sci.* **2**, 71–87 (1991).
5. Barrett, H. C. & Fiddick, L. Evolution and risky decisions. *Trends Cogn. Sci.* **4**, 251–252 (2000).
6. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus–norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
7. Harvey, S. C. Non-dauer larval dispersal in *Caenorhabditis elegans*. *J. Exp. Zool. B* **312**, 224–230 (2009).
8. Shtonda, B. B. & Avery, L. Dietary choice behavior in *Caenorhabditis elegans*. *J. Exp. Biol.* **209**, 89–102 (2006).
9. Pujol, N. et al. A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Curr. Biol.* **11**, 809–821 (2001).
10. Pradel, E. et al. Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **104**, 2295–2300 (2007).
11. Lipton, J., Kleemann, G., Ghosh, R., Lints, R. & Emmons, S. W. Mate searching in *Caenorhabditis elegans*: a genetic model for sex drive in a simple invertebrate. *J. Neurosci.* **24**, 7427–7434 (2004).
12. Gloria-Soria, A. & Azevedo, R. B. *npr-1* Regulates foraging and dispersal strategies in *Caenorhabditis elegans*. *Curr. Biol.* **18**, 1694–1699 (2008).
13. Styer, K. L. et al. Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science* **322**, 460–464 (2008).
14. Reddy, K. C., Andersen, E. C., Kruglyak, L. & Kim, D. H. A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science* **323**, 382–384 (2009).
15. de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**, 679–689 (1998).
16. Gray, J. M. et al. Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. *Nature* **430**, 317–322 (2004).
17. Rogers, C., Persson, A., Cheung, B. & de Bono, M. Behavioral motifs and neural pathways coordinating O₂ responses and aggregation in *C. elegans*. *Curr. Biol.* **16**, 649–659 (2006).
18. Bretscher, A. J., Busch, K. E. & de Bono, M. A carbon dioxide avoidance behavior is integrated with responses to ambient oxygen and food in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **105**, 8044–8049 (2008).
19. McGrath, P. T. et al. Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* **61**, 692–699 (2009).
20. Macosko, E. Z. et al. A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* **458**, 1171–1175 (2009).
21. Osborne, K. A. et al. Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science* **277**, 834–836 (1997).
22. Reaume, C. J. & Sokolowski, M. B. cGMP-dependent protein kinase as a modifier of behaviour. *Handb. Exp. Pharmacol.* **191**, 423–443 (2009).
23. Hong, R. L., Witte, H. & Sommer, R. J. Natural variation in *Pristionchus pacificus* insect pheromone attraction involves the protein kinase EGL-4. *Proc. Natl Acad. Sci. USA* **105**, 7779–7784 (2008).
24. Rockman, M. V. & Kruglyak, L. Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet.* **5**, e1000419 (2009).
25. Steinmetz, L. M. et al. Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**, 326–330 (2002).
26. Edwards, A. C. & Mackay, T. F. Quantitative trait loci for aggressive behavior in *Drosophila melanogaster*. *Genetics* **182**, 889–897 (2009).

27. Legare, M. E., Bartlett, F. S. II & Frankel, W. N. A major effect QTL determined by multiple genes in epileptic EL mice. *Genome Res.* **10**, 42–48 (2000).
28. Thomson, M. J., Edwards, J. D., Septiningsih, E. M., Harrington, S. E. & McCouch, S. R. Substitution mapping of *dth1.1*, a flowering-time quantitative trait locus (QTL) associated with transgressive variation in rice, reveals multiple sub-QTL. *Genetics* **172**, 2501–2514 (2006).
29. Wragg, R. T. *et al.* Tyramine and octopamine independently inhibit serotonin-stimulated aversive behaviors in *Caenorhabditis elegans* through two novel amine receptors. *J. Neurosci.* **27**, 13402–13412 (2007).
30. Mackay, T. F. Quantitative trait loci in *Drosophila*. *Nature Rev. Genet.* **2**, 11–20 (2001).
31. Bargmann, C. I. & Horvitz, H. R. Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in *C. elegans*. *Neuron* **7**, 729–742 (1991).
32. Gray, J. M., Hill, J. J. & Bargmann, C. I. A circuit for navigation in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **102**, 3184–3191 (2005).
33. Wakabayashi, T., Kitagawa, I. & Shingai, R. Neurons regulating the duration of forward locomotion in *Caenorhabditis elegans*. *Neurosci. Res.* **50**, 103–111 (2004).
34. Hallem, E. A. & Sternberg, P. W. Acute carbon dioxide avoidance in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **105**, 8038–8043 (2008).
35. Zimmer, M. *et al.* Neurons detect increases and decreases in oxygen levels using distinct guanylate cyclases. *Neuron* **61**, 865–879 (2009).
36. Frøkjær-Jensen, C. *et al.* Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nature Genet.* **40**, 1375–1383 (2008).
37. Suo, S., Kimura, Y. & Van Tol, H. H. Starvation induces cAMP response element-binding protein-dependent gene expression through octopamine-Gq signaling in *Caenorhabditis elegans*. *J. Neurosci.* **26**, 10082–10090 (2006).
38. Alkema, M. J., Hunter-Ensor, M., Ringstad, N. & Horvitz, H. R. Tyramine functions independently of octopamine in the *Caenorhabditis elegans* nervous system. *Neuron* **46**, 247–260 (2005).
39. Greer, E. R., Perez, C. L., Van Gilst, M. R., Lee, B. H. & Ashrafi, K. Neural and molecular dissection of a *C. elegans* sensory circuit that regulates fat and feeding. *Cell Metab.* **8**, 118–131 (2008).
40. Pirri, J. K., McPherson, A. D., Donnelly, J. L., Francis, M. M. & Alkema, M. J. A tyramine-gated chloride channel coordinates distinct motor programs of a *Caenorhabditis elegans* escape response. *Neuron* **62**, 526–538 (2009).
41. Crocker, A. & Sehgal, A. Octopamine regulates sleep in *Drosophila* through protein kinase A-dependent mechanisms. *J. Neurosci.* **28**, 9377–9385 (2008).
42. Roeder, T. Tyramine and octopamine: ruling behavior and metabolism. *Annu. Rev. Entomol.* **50**, 447–477 (2005).
43. Hoyer, S. C. *et al.* Octopamine in male aggression of *Drosophila*. *Curr. Biol.* **18**, 159–167 (2008).
44. Yalcin, B. *et al.* Genetic dissection of a behavioral quantitative trait locus shows that *Rgs2* modulates anxiety in mice. *Nature Genet.* **36**, 1197–1202 (2004).
45. Young, L. J., Nilsen, R., Waymire, K. G., MacGregor, G. R. & Insel, T. R. Increased affiliative response to vasopressin in mice expressing the V_{1a} receptor from a monogamous vole. *Nature* **400**, 766–768 (1999).
46. Mackay, T. F., Stone, E. A. & Ayroles, J. F. The genetics of quantitative traits: challenges and prospects. *Nature Rev. Genet.* **10**, 565–577 (2009).
47. Gerke, J., Lorenz, K., Ramnarine, S. & Cohen, B. Gene-environment interactions at nucleotide resolution. *PLoS Genet.* **6**, e1001144 (2010).
48. Ramot, D., Johnson, B. E., Berry, T. L. Jr, Carnell, L. & Goodman, M. B. The Parallel Worm Tracker: a platform for measuring average speed and drug-induced paralysis in nematodes. *PLoS ONE* **3**, e2208 (2008).
49. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890 (2003).
50. Mello, C. & Fire, A. DNA transformation. *Methods Cell Biol.* **48**, 451–482 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Shingai for strains, and P. McGrath and members of the C.I.B. laboratory for discussions. A.B. was supported by the Secretaría de Educación Pública of Mexico and by The Rockefeller University. C.I.B. and L.K. are Investigators of the Howard Hughes Medical Institute. This work was supported by HHMI and by NIH grant GM089972.

Author Contributions A.B. and C.I.B. designed experiments, A.B. conducted experiments, M.V.R. constructed strains for QTL mapping, M.T. developed tracking methods, A.B., M.V.R., L.K. and C.I.B. analysed and interpreted results, and A.B. and C.I.B. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.I.B. (cori@rockefeller.edu).

METHODS

Analysis of behaviour in the leaving assay. Six-centimetre NGM agar plates were seeded with 70 μ l (conditioning plate) or with 10 μ l (assay plate) of a fresh overnight culture of *Escherichia coli* HB101 diluted in LB to $A_{600nm} = 2.0$. 90 min after seeding the plates, ten young adult hermaphrodites were picked onto the conditioning plate. 30 min after being placed on the conditioning plates, seven of the animals were transferred onto the lawn of the assay plate. The 30 min leaving assay began 1 h after placing the seven animals on the assay plate. The number of leaving events was recorded manually by examining the video recordings, and further behavioural analysis was conducted with a Matlab code adapted from the Parallel Worm Tracker⁴⁸. A leaving event was defined as an episode in which the whole body of an animal left the bacterial lawn and the animal did not reverse immediately to return to the lawn. The leaving rate was calculated as the number of leaving events per worm minute spent inside the bacterial lawn. Experiments on each strain were repeated at least three times.

Quantitative trait locus analysis. The N2-HW recombinant inbred advanced intercross lines (RIAILs) used in this study represent the terminal generation of a 20-generation pedigree founded by reciprocal crosses between N2 and HW. The lines were constructed through 10 generations of intercrossing followed by 10 generations of selfing²⁴. They have been genotyped at 1,454 nuclear and one mitochondrial markers and have a 5.3-fold expansion of the F_2 genetic map²⁴. QTL analysis was performed on the mean leaving rates of N2-HW RIAILs by non-parametric interval mapping in R/qtl⁴⁹. Significance levels were estimated from 10,000 permutations of the data.

Quantitative RT-PCR. Total RNA from mixed stage worms was isolated with Trizol. 1.5 μ g of RNA and oligo-dT were used for reverse transcription using SuperScript III First-Strand Synthesis (Invitrogen) according to the manufacturer's instructions. Real-time PCR was performed with Fast SYBR Green Master Mix (Applied Biosystems) on a 7900HT Real-Time PCR System (Applied Biosystems). *act-3* was used as the calibrator for relative quantitation. 5' primers corresponded to upstream exons that distinguished *tyra-3* isoforms, and 3' primers corresponded to shared exon sequence. Primers used were: *tyra-3a* & *tyra-3c.2_F*, ccacttgcaaatagcagcag; *tyra-3b_F*, ggctatttggtggtgttg; *tyra-3a* & *tyra-3b_R*, tcctctggcgctcgaaatag; *act-3_F*, tcacatcatgagaccattcaaa; *act-3_R*, gcaaatgtagtgggtcttcttatg.

***tyra-3* expression pattern.** The N2 and HW 4.9 kb *tyra-3b* promoters were amplified using primers: tcaacctaaccactaactaagg and cGatgaagcaagatgtcaggt, which overlaps the coding region by 4 bp. The ATG start codon is mutated to ATC (mutation is uppercase in primer). These promoters were individually fused by PCR to a fragment containing GFP followed by the *unc-54* 3'-UTR, as described⁵¹. These PCR products were injected individually into both HW and N2 animals at 20 ng μ l⁻¹. Cells expressing GFP were identified by Nomarski microscopy in both L1 and adult hermaphrodites. The identification of some cells was aided by injecting *Ptyra-3b::GFP*-expressing animals with promoter-mCherry fusions with established expression patterns. In this manner, the AIM neurons were identified as *Ptyra-3b::GFP*-expressing cells based on their position and the absence of co-localization with *Pttx-3::mCherry*. The BAG neurons co-expressed *Ptyra-3b::GFP* and *Pflp-17::mCherry*. The CEP neurons co-expressed *Ptyra-3b::GFP* and *Pdat-1::mCherry*. The ASK neurons co-expressed *Ptyra-3b::GFP* and *Psra-9::mCherry*. The ADL neurons co-expressed *Ptyra-3b::GFP* and *Psri-51::mCherry*.

Extrachromosomal transgenes. Transgenes were made by injection of DNA clones into the gonads of young adult hermaphrodites together with a fluorescent coinjection marker⁵⁰. To control for variation between transgenes, two to five independent lines from each injection were characterized.

Generation of MosSCI lines and quantitation of GFP fluorescence in ASK. Single-copy insertion of transgenes was performed using the direct MosSCI

technique targeting the *ttTi5605* Mos allele on chromosome II, as described³⁶. A schematic of the mechanism underlying MosSCI is shown in Fig. 4c.

The pCFJ151 targeting vector was modified by the introduction of an FseI restriction site into the multiple cloning site by site-directed mutagenesis using the primers gtaatacgaactcacttaagccggccctagagggtaccagagctcacc and ggtgagctctggt acccttagggccggccttaagtgcgtgtattac to make pAB1. An FseI-SpeI fragment from a pSM vector containing *N2-Ptyra-3b::N2-tyra-3b::SL2 GFP::unc-54* 3'-UTR or *HW-Ptyra-3b::N2-tyra-3b::SL2 GFP::unc-54* 3'-UTR was cloned into pAB1.

For each *tyra-3*-containing test plasmid, about 50 EG4322 animals were injected with a mixture of *tyra-3* plasmid, pGH8, pCFJ90, pCFJ104 and pJL43.1. After positive and negative selection and full sequencing of the insert, two inserted transgenes each of *N2-Ptyra-3b* and *HW-Ptyra-3b* were backcrossed to HW males seven times, selecting GFP-fluorescent hermaphrodites each generation. The transgene-containing chromosome was then homozygosed.

The strains containing the single-copy transgene in a HW background were injected with *Psra-9::mCherry* to identify ASK. Young adult hermaphrodites were examined on a Zeiss Imager Z.1 with a $\times 60$ objective focused on ASK using mCherry to prevent bleaching of GFP signal. Fluorescence signals were acquired with fixed acquisition times (30–50 ms for mCherry, 100 ms for GFP). Background mean fluorescence intensity adjacent to ASK was subtracted from the ASK signal.

RNAi. RNA interference was performed essentially as described⁵². A fragment common to all *tyra-3* isoforms was amplified. The following primers were used, which include the T7 sequence (underlined): taatacgaactcactataggagagaaatggc agcaggactctt; taatacgaactcactataggagaatcctgcagctgtggagt. *in vitro* transcription was performed with RiboMAX kit (Promega). dsRNA was injected at 1.2 μ g μ l⁻¹ into the gonads of adult hermaphrodites. Eggs laid 24 and 48 h after injection were used for the behavioural assays.

Octanol avoidance assay. Avoidance assays were conducted essentially as described⁵³. In brief, ~20 three-day-old animals were picked off of their growth plates food into a transfer plate without bacteria where they were allowed to crawl and rid themselves of bacteria. Animals were then transferred onto an NGM plate without food. After 40 min, a microcapillary with 30% octanol (v/v diluted fresh every day in ethanol) was presented in front of the animal's nose. The time to reverse was recorded. If animals did not reverse within 20 s, the assay was stopped. Animals were presented with odour 1–3 times per experiment, with at least 3 min of rest interval. We replicated published results demonstrating that *tyra-3* null mutants had more rapid responses than N2 in the presence of exogenous serotonin and tyramine²⁹ but also observed more rapid responses in the absence of exogenous neuromodulators, as shown in Supplementary Fig. 6.

Cell ablations. For leaving behaviour assays, ASK was ablated with a laser microbeam as described⁵⁴. BAG was killed using split human caspase 3 fragments⁵⁵ expressed from *flp-17* and *glb-5* promoters that overlapped only in BAG. For lysine chemotaxis assays, ASK was killed using a mouse caspase 1 gene expressed from the *sra-9* promoter⁵⁶. The ASK strain was a gift from R. Shingai.

- Hobert, O. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechnology* **32**, 728–730 (2002).
- Ahringer, J. Reverse genetics. In *WormBook* (ed. The *C. elegans* Research Community) doi:10.1895/wormbook.1.47.1 (6 April 2006); available at (<http://www.wormbook.org>).
- Troemel, E. R., Chou, J. H., Dwyer, N. D., Colbert, H. A. & Bargmann, C. I. Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*. *Cell* **83**, 207–218 (1995).
- Bargmann, C. I. & Avery, L. Laser killing of cells in *Caenorhabditis elegans*. *Methods Cell Biol.* **48**, 225–250 (1995).
- Chelur, D. S. & Chalfie, M. Targeted cell killing by reconstituted caspases. *Proc. Natl Acad. Sci. USA* **104**, 2283–2288 (2007).
- Kim, K. et al. Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. *Science* **326**, 994–998 (2009).

Collapse of long-range charge order tracked by time-resolved photoemission at high momenta

Timm Rohwer^{1*}, Stefan Hellmann^{1*}, Martin Wiesenmayer¹, Christian Sohrt¹, Ankatrin Stange¹, Bartosz Slomski¹, Adra Carr², Yanwei Liu^{3,4}, Luis Mijang Avila⁵, Matthias Kalläne¹, Stefan Mathias^{2,6}, Lutz Kipp¹, Kai Rossnagel¹ & Michael Bauer¹

Intense femtosecond (10^{-15} s) light pulses can be used to transform electronic, magnetic and structural order in condensed-matter systems on timescales of electronic and atomic motion^{1,2,3}. This technique is particularly useful in the study^{4,5} and in the control⁶ of materials whose physical properties are governed by the interactions between multiple degrees of freedom. Time- and angle-resolved photoemission spectroscopy is in this context a direct and comprehensive, energy- and momentum-selective probe of the ultrafast processes that couple to the electronic degrees of freedom^{7–10}. Previously, the capability of such studies to access electron momentum space away from zero momentum was, however, restricted owing to limitations of the available probing photon energy^{10,11}. Here, using femtosecond extreme-ultraviolet pulses delivered by a high-harmonic-generation source, we use time- and angle-resolved photoemission spectroscopy to measure the photoinduced vaporization of a charge-ordered state in the potential excitonic insulator 1T-TiSe₂ (refs 12, 13). By way of stroboscopic imaging of electronic band dispersions at large momentum, in the vicinity of the edge of the first Brillouin zone, we reveal that the collapse of atomic-scale periodic long-range order happens on a timescale as short as 20 femtoseconds. The surprisingly fast response of the system is assigned to screening by the transient generation of free charge carriers. Similar screening scenarios are likely to be relevant in other photoinduced solid-state transitions and may generally determine the response times. Moreover, as electron states with large momenta govern fundamental electronic properties in condensed matter systems¹⁴, we anticipate that the experimental advance represented by the present study will be useful to study the ultrafast dynamics and microscopic mechanisms of electronic phenomena in a wide range of materials.

The electronic properties of many condensed matter systems are determined by large-momentum electron states, often located near the edge of the first Brillouin zone (BZ), the unit cell of crystalline solids in electron momentum (k) space. A prominent current example is graphene, whose hallmark, the critical crossing point of its peculiar conical band dispersion, is located at the corner of the BZ¹⁵. The copper oxide based high-temperature superconductors represent another example, where the much debated competition between the pseudogap and superconductivity is particularly pronounced in the so-called antinodal region of the Fermi surface, near the BZ boundary¹⁶. A final example is the new class of iron-based superconductors, which are characterized by magnetic excitations that couple two sets of Fermi surfaces, one set centred on the corners of the BZ¹⁷. More generally, it is the coupling of high momentum electron states near the Fermi momenta k_F (the momenta of the highest occupied electron states) that contributes most to the linear response function of an electron liquid, sometimes even causing a divergence that leads to phase instabilities.

Typical electron momenta near the boundary of the first BZ are in the 1 \AA^{-1} regime. Conventional angle-resolved photoemission spectroscopy (ARPES) with photon energies exceeding roughly 10 eV is probably the most powerful tool to map band structure peculiarities and Fermi surfaces up to and beyond these critical points: this technique can also, at the same time, determine many-body effects embodied in the fine details of band dispersions and in the distribution of spectral weight. ARPES is particularly well suited for layered materials—as in the present study—because for quasi-two-dimensional systems, the measured electronic structure can be considered as being predominantly characteristic of the bulk, despite the surface sensitivity of the probe. The great allure of corresponding time-resolved ARPES experiments is the provision of direct dynamical information and the possibility of disentangling—via temporal discrimination—the various interactions between the relevant degrees of freedom that determine material properties in the quantum world. Here, we present femtosecond time-resolved ARPES experiments, in which transient changes in the whole occupied electronic structure between the centre and the edge of the BZ are probed to answer a simple fundamental question: how fast can long-range charge order in a solid melt?

The charge-ordered state we investigate is the conspicuous ($2 \times 2 \times 2$) charge-density wave (CDW) that occurs in the layered compound 1T-TiSe₂ below a temperature of 200 K (ref. 13). Figure 1 presents the thermal equilibrium view of the CDW transition, which affects both structural and electronic properties. On the transition, the atoms move to new equilibrium positions such that the real-space unit cell doubles its size in all three directions (Fig. 1a and c; arrows in Fig. 1c indicate the atomic displacements from the normal-phase positions); in momentum space, the dimensions of the BZ are correspondingly halved (Fig. 1b and d). The new k -space geometry suggests that the wave vector of the CDW is determined by an interaction between the Se 4p valence band maximum at $\bar{\Gamma}$, the centre of the BZ, and the elliptical pocket of Ti 3d states at the BZ edge at \bar{M} (Fig. 1b). These symmetry points are connected by the new reciprocal lattice vectors and become equivalent $\bar{\Gamma}$ points in the new phase, which allows direct Se 4p–Ti 3d interaction (Fig. 1d). This interaction is in fact remarkably strong and extremely well resolved by conventional ARPES^{12,18,19}. Figure 1e and f compares ARPES intensity maps recorded with synchrotron radiation ($h\nu = 119 \text{ eV}$) along the $\bar{\Gamma}$ – \bar{M} direction above and below the CDW transition temperature. At $\bar{\Gamma}$, the downward dispersing (hole-like) Se 4p bands dominate the photoemission signal in both maps and only a small shift of the valence band maximum is visible. At \bar{M} , where the high-temperature map shows the bottom of the upward dispersing (electron-like) Ti 3d band, the effects are more dramatic: the CDW leads to a strong selective transfer of spectral weight by folding the Se 4p band from $\bar{\Gamma}$ onto \bar{M} . It is this remarkably strong folded Se 4p intensity which we will use in the time-resolved experiments as a spectroscopic measure for CDW order in 1T-TiSe₂.

¹Institute of Experimental and Applied Physics, University of Kiel, 24118 Kiel, Germany. ²JILA, University of Colorado and NIST, Boulder, Colorado 80309-0440, USA. ³University of California, Berkeley, NSF ERC Extreme Ultraviolet Science and Technology, Berkeley, California 94720, USA. ⁴Center for X-Ray Optics, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁵Center for Nano and Molecular Science, University of Texas at Austin, Austin, Texas 78722, USA. ⁶Department of Physics and Research Center OPTIMAS, University of Kaiserslautern, 67663 Kaiserslautern, Germany.

*These authors contributed equally to this work.

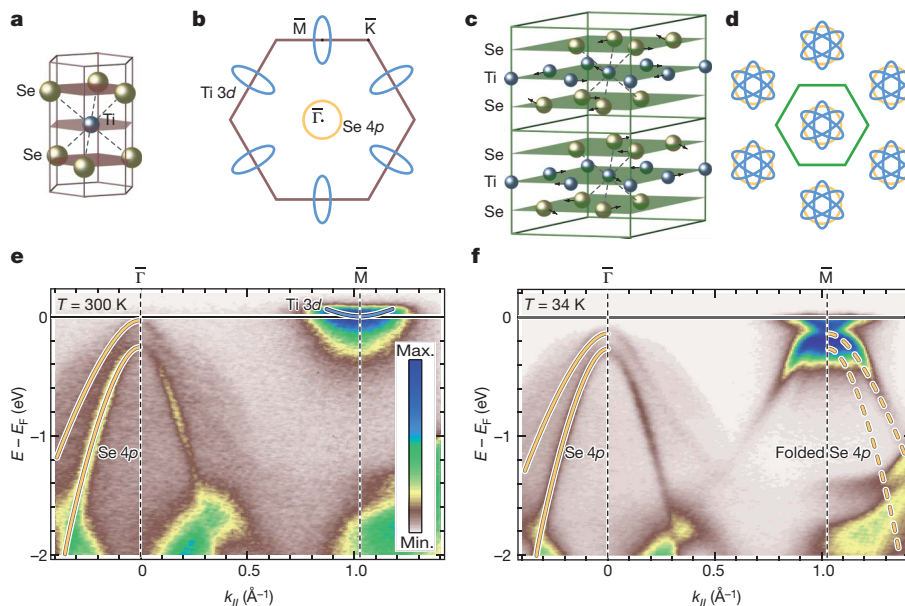


Figure 1 | CDW phase transition of 1T-TiSe₂. **a**, Real-space unit cell of the normal phase at room temperature. **b**, Common unit cell (first BZ, purple line) in momentum space of the room temperature phase (two-dimensional projection onto the surface plane). $\bar{\Gamma}$, \bar{M} and \bar{K} are high-symmetry points of the first BZ. Blue ellipses and yellow circle indicate the Fermi surface topology of Ti 3d and Se 4p bands, respectively. Planes show the atomic layers of the two-dimensional crystalline structure. **c**, Real-space unit cell of the CDW phase.

We now apply sub-10-fs extreme ultraviolet (XUV) pulses ($h\nu = 43$ eV, s-polarization) to monitor the transient response of the CDW phase to excitation with infrared laser pulses ($h\nu = 1.57$ eV) of 32 fs width. Details of the time-resolved ARPES experiment are described in Methods. Figure 2a shows ARPES intensity maps of 1T-TiSe₂ measured at $T = 125$ K with the femtosecond XUV light source²⁰ (Methods). Despite the poorer energy resolution, both sets of Se 4p bands are well resolved, the original one at $\bar{\Gamma}$ and the folded one at \bar{M} . Time-resolved experiments are performed at infrared pump fluences between 0.2 and 5 mJ cm⁻², corresponding to an excitation density range of 0.025 to 0.63 photons per Ti atom. Figure 2b–e shows four photoemission snapshots recorded at a pump fluence of 5 mJ cm⁻² with increasing temporal delay between the infrared pump and the XUV probe, up to a maximum of 3 ps (Supplementary Movie 1). The data have been corrected for a space charge shift of 200 meV induced by the electron background because of multi-photon photoemission by the infrared pump pulse. The time series is dominated by two prominent changes in the photoemission intensity maps. First, in instantaneous response to the infrared excitation, an electron-like band appears, crossing the Fermi energy E_F and extending (at sufficiently small temporal delays) from \bar{M} to $\bar{\Gamma}$. We observe here the transient generation of quasi-free charge carriers because of near-resonant Ti 3d – Se 4p excitation. Second, the downward-dispersing Se 4p band, folded onto the \bar{M} point owing to the interaction with the CDW superlattice, disappears or is at least considerably reduced in intensity. This suggests that long-range order in the electronic subsystem breaks down on an ultrafast timescale. In the following, we restrict our quantitative analysis to the short-time (sub-100-fs) dynamics of this process.

Figure 2f compares the temporal evolution of the integrated intensity of the folded Se 4p band—our spectroscopic measure for CDW order—for different pump fluences (see Supplementary Information section 3 for details of the data analysis). Both breakdown and (partial) recovery of the signal (inset of Fig. 2f) are strongly dependent on the pump fluence. The fluence dependence of the time constant characterising the signal breakdown, $\tau_{\text{Se } 4p}$, is shown in Fig. 3a: at the lowest fluences,

Arrows indicate the atomic displacements from the normal-phase positions; the dashed lines indicate the extension of the unit cell in the normal phase.

d, First BZ (green line) of the CDW phase. The folding of Se 4p and Ti 3d states is indicated. **e**, ARPES intensity map (electron binding energy versus momentum) of the room temperature phase. Photoelectron intensity is encoded in a false-colour scale. **f**, ARPES intensity map of the CDW phase.

the initial drop in the signal is retarded by about 80 fs with respect to the laser pulse excitation. As the fluence increases, the response becomes continuously faster, and at the highest fluences the transient minimum in the folded Se 4p band intensity appears well within the 32-fs-long infrared pump laser pulse with an ultimate response time of 20 fs. For comparison, the dynamics associated with the initial population of the Ti 3d band due to absorption happens within the width of the infrared pulse for the entire pump fluence regime. Notably, for the highest excitation fluence, the folded Se 4p band follows this population dynamics without delay (Supplementary Information section 4). The partial recovery of the folded Se 4p intensity is observed on timescales of several hundreds of femtoseconds. Two-temperature model calculations following reference 21 suggest that this recovery is mostly driven by thermalization of the electronic subsystem with the atomic lattice.

In previous studies, it has been shown that the fundamental timescales of photoinduced phase transitions are governed by bottlenecks associated with the characteristic response times of the relevant degrees of freedom, such as the oscillation period of neighbouring atoms²² or the hopping rate of localized electrons between neighbouring sites²³. The upper solid line in Fig. 3a marks for instance the expected short-time limit (75 fs) of the lattice response of 1T-TiSe₂ to a photoexcitation; this short-time limit is taken as one-quarter of the oscillation period of the high-frequency CDW amplitude mode²⁴. The vaporization of a long-range-ordered state within 20 fs is in this context exceptionally fast. The ultrafast timescale observed in the high-fluence regime points to a purely electronically driven process, whose response time, however, strongly depends on the excitation fluence. The absorption of the light pulse initially increases the free charge carrier density n (electrons and holes), as can be seen in the instantaneous population of the Ti 3d band. This transient free carrier population, which is directly governed by the excitation fluence, links the time constant $\tau_{\text{Se } 4p}$ to a material specific timescale: quantum kinetic calculations have shown that the characteristic build-up time for carrier screening in response to an ultrashort laser excitation is the plasma oscillation period, τ_{pl} (refs 25, 26), which scales with $1/\sqrt{n}$. As shown in Fig. 3a, $\tau_{\text{Se } 4p}$ closely follows such a $1/\sqrt{n}$ dependence (a quantitative estimate of the

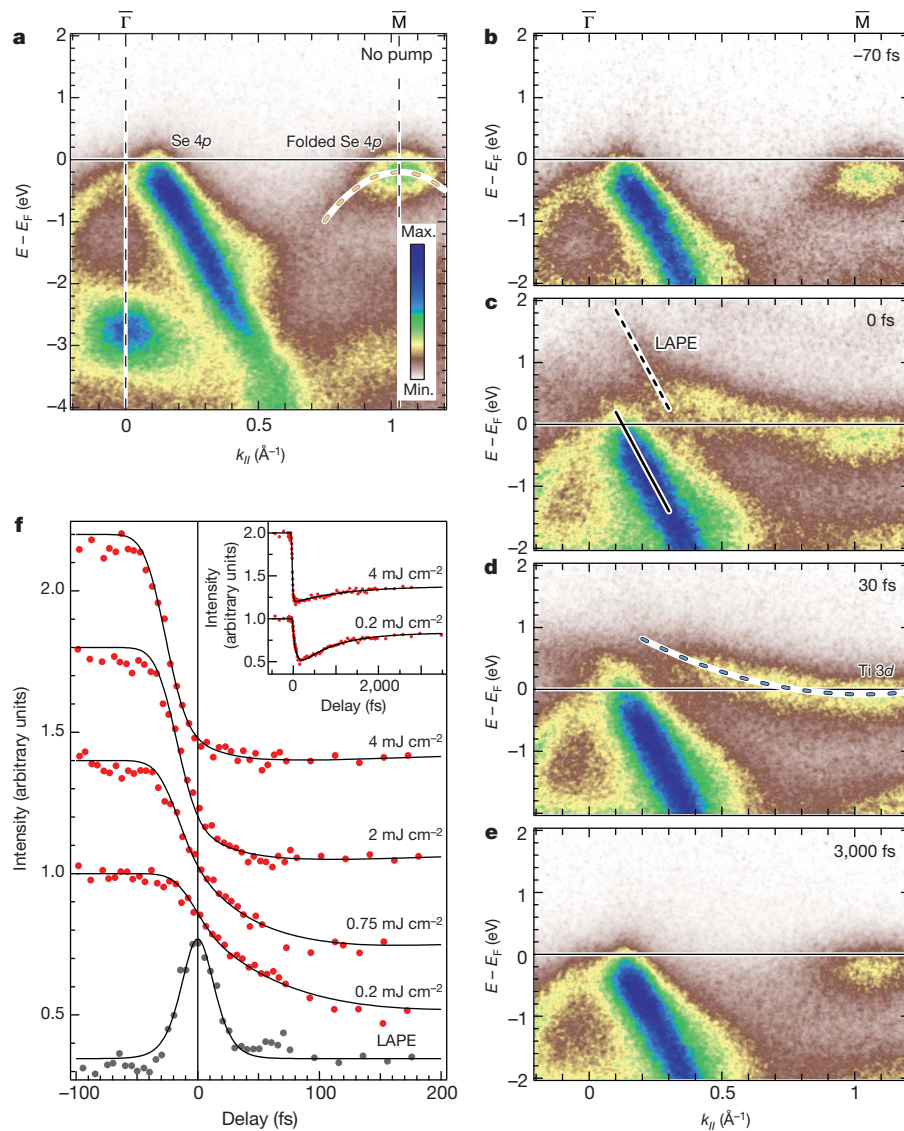


Figure 2 | Tracking the photoinduced transition by femtosecond time-resolved ARPES. **a**, ARPES intensity map of 1T-TiSe₂ recorded with high-harmonic XUV pulses in the CDW phase (temperature $T = 125$ K). **b–e**, Time-resolved ARPES snapshots at increasing pump (infrared)–probe (XUV) temporal delays (pump–pulse fluence, 5 mJ cm^{-2}). Energy distribution curves (EDCs) are provided in Supplementary Information 2. **f**, Photoemission transients of the folded Se 4p band for different infrared excitation fluences close to time zero. The infrared–XUV cross-correlation signal, which has been determined from the laser assisted photoemission³⁰ (LAPE) signal, is added (Methods). Inset, transients up to temporal delays of 3.5 ps.

photoinduced carrier density n and the corresponding plasma oscillation period is given in Supplementary Information section 5). It has in fact been shown that the build-up of screening by photo-injected carriers is relevant for the dynamics of ultrafast processes on the sub-100-fs timescale²⁷.

Suppression of screening is important in correlation-induced phase transitions. Therefore, it is not surprising that the reverse—the enhancement of screening by photo-injection of free carriers—can destroy a correlation-induced phase. The CDW transition in 1T-TiSe₂ has repeatedly been associated with an excitonic insulator instability^{12,13}. In this model, the transition is driven by the spontaneous formation of excitons, which can occur in semiconductors or semimetals when the bandgap or band overlap becomes smaller than the exciton binding energy. The exciton formation requires a low concentration of mobile charge carriers and a correspondingly poorly screened Coulomb interaction. In the case of 1T-TiSe₂, the narrow gap/overlap system would become unstable to the formation of Ti 3d – Se 4p excitons and would exhibit a new periodicity governed by the wave vector connecting the corresponding valence and conduction band pockets. This purely electronic process seems consistent with the measured ultrafast response times and a screening-based interpretation, as bound excitons would certainly be screened by the photo-injected carriers. For a transition temperature of 200 K, energy-time uncertainty yields a response time of 35 fs for such a purely excitonic process (see lower solid line in Fig. 3a). More generally, however, our results are in line with the

screening of any (unspecified) interaction between Se 4p and Ti 3d states described by an effective matrix element V_{eff}^2 . The spectral weight of the folded Se 4p state will scale with V_{eff}^2 and we can estimate the effect of screening within the Thomas-Fermi approach¹⁴. This yields $V_{\text{eff}}^2 \approx (1 + \text{const.} \times n^{1/3})$, in reasonable agreement with our experimental data (see Fig. 3b). A deeper analysis, particularly with respect to the puzzling story of the CDW phase transition in 1T-TiSe₂, requires a sophisticated theoretical description which, for instance, considers the quantum kinetics of screening in non-equilibrium systems or effects arising from the photo-doping of excitonic insulators.

We finally address the question of to what extent the screening by the nascent carriers affects the properties of 1T-TiSe₂ on the 20-fs timescale. As discussed above, the response of the atomic lattice is slow so that structurally the sample will still resemble the CDW phase on this timescale. However, the valence electronic structure becomes substantially modified as soon as screening becomes effective. In fact, electronically, the system undergoes a transition from a poorly conducting CDW state into a metallic phase within 20 fs. This is not only implied by the screening scenario but is also directly visible in Fig. 2d, which shows the transient metallization of the Ti 3d band only 30 fs after the optical excitation.

The ultrafast breakdown of long-range charge order that we report here is much faster than the material-characteristic oscillations of collective modes that are commonly thought to limit the response times in photoinduced processes. This surprising result may therefore stimulate

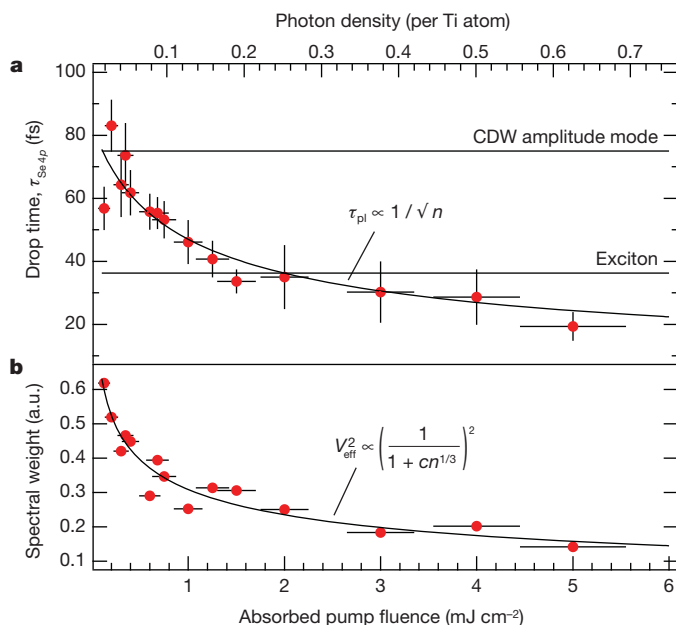


Figure 3 | Fluence dependence of the photoinduced transition. **a**, Signal drop time, $\tau_{\text{Se } 4p}$, of the folded Se 4p band (filled red circles) as a function of the absorbed pump fluence. The $\sim 10\%$ error bars in the absorbed pump fluence are determined from the uncertainty in the optical constants of 1T-TiSe₂, the stability of the pump laser, and the uncertainty in the pump pulse diameter at the sample position. The error bars in the drop time indicate the standard deviation of the fits to the data. **b**, Minimum transient spectral weight in the folded Se 4p band (red dots) as a function of absorbed fluence. Equation of curve fitted to data is shown in each panel. a.u., arbitrary units.

new concepts for ultrafast switching devices. Furthermore, in our study we have essentially monitored the intensity of a superlattice Bragg peak in electron momentum space. Thus, from a methodological point of view, the presented time-resolved ARPES approach complements present-day time-resolved diffraction experiments^{28,29}, with the advantage of an exceptionally high temporal resolution.

METHODS SUMMARY

Sample preparation. 1T-TiSe₂ single crystals were grown from the elements by chemical vapour transport using iodine as transport agent. Before the photoemission measurements, the samples were cleaved *in situ* at room temperature in ultrahigh vacuum.

Photoemission measurements. Static ARPES experiments were conducted at beamline 7.0.1 of the Advanced Light Source at Berkeley with a Scienta R4000 electron spectrometer. The photon energy was 119 eV and the overall energy resolution was ~ 30 meV. Femtosecond time-resolved ARPES measurements were conducted at the University of Kiel with a SPECS Phoibos 150 electron spectrometer. Here, the photon energy was 43 eV and the overall energy resolution in the experiment was ~ 400 meV. The light source for the time-resolved experiments was an argon-filled hollow-fibre waveguide (XUUS, KMLabs) for high harmonic generation operated with a 3 kHz Ti:sapphire amplifier system (Dragon, KMLabs, pumped by an Empower 30, Spectra Physics).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 September 2010; accepted 11 January 2011.

Published online 9 March 2011.

- Yonemitsu, K. & Nasu, K. Theory of photoinduced phase transitions in itinerant electron systems. *Phys. Rep.* **465**, 1–60 (2008).
- Bargheer, M., Zhavoronkov, N., Woerner, M. & Elsaesser, T. Recent progress in ultrafast X-ray diffraction. *ChemPhysChem* **7**, 783–792 (2006).
- Bigot, J.-Y., Vomir, M. & Beaupaire, E. Coherent ultrafast magnetism induced by femtosecond laser pulses. *Nature Phys.* **5**, 515–520 (2009).

- Chollet, M. *et al.* Gigantic photoresponse in $\frac{1}{4}$ -filled-band organic salt (EDO-TTF)₂PF₆. *Science* **307**, 86–89 (2005).
- Demsar, J., Biljaković, K. & Mihailović, D. Single particle and collective excitations in the one-dimensional charge density wave solid K_{0.3}MoO₃ probed in real time by femtosecond spectroscopy. *Phys. Rev. Lett.* **83**, 800–803 (1999).
- Rini, M. *et al.* Control of the electronic phase of a manganite by mode-selective vibrational excitation. *Nature* **449**, 72–74 (2007).
- Haight, R. & Silberman, J. A. Surface intervalley scattering on GaAs(110): direct observation with picosecond laser photoemission. *Phys. Rev. Lett.* **62**, 815–818 (1989).
- Fann, W. S., Storz, R., Tom, H. W. K. & Bokor, J. Electron thermalization in gold. *Phys. Rev. B* **46**, 13592–13595 (1992).
- Bauer, M. *et al.* Direct observation of surface chemistry using ultrafast soft-X-ray pulses. *Phys. Rev. Lett.* **87**, 025501 (2001).
- Perfetti, L. *et al.* Time evolution of the electronic structure of 1T-TaS₂ through the insulator-metal transition. *Phys. Rev. Lett.* **97**, 067402 (2006).
- Schmitt, F. *et al.* Transient electronic structure and melting of a charge density wave in TbTe₃. *Science* **321**, 1649–1652 (2008).
- Cercellier, H. *et al.* Evidence for an excitonic insulator phase in 1T-TiSe₂. *Phys. Rev. Lett.* **99**, 146403 (2007).
- Di Salvo, F. J., Moncton, D. E. & Waszczak, J. V. Electronic properties and superlattice formation in the semimetal TiSe₂. *Phys. Rev. B* **14**, 4321–4328 (1976).
- Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* (Brooks/Cole, Belmont, 1976).
- Ohta, T., Bostwick, A., Seyller, T., Horn, K. & Rotenberg, E. Controlling the electronic structure of bilayer graphene. *Science* **313**, 951–954 (2006).
- Kondo, T., Khasanov, R., Takeuchi, T., Schmalian, J. & Kaminski, A. Competition between the pseudogap and superconductivity in the high-*T_c* copper oxides. *Nature* **457**, 296–300 (2009).
- Mazin, I. Superconductivity gets an iron boost. *Nature* **464**, 183–186 (2010).
- Rosnagel, K., Kipp, L. & Skibowski, M. Charge-density-wave phase transition in 1T-TiSe₂: excitonic insulator versus band-type Jahn-Teller mechanism. *Phys. Rev. B* **65**, 235101 (2002).
- Kidd, T. E., Miller, T., Chou, M. Y. & Chiang, T.-C. Electron-hole coupling and the charge density wave transition in TiSe₂. *Phys. Rev. Lett.* **88**, 226402 (2002).
- Rundquist, A. *et al.* Phase matching of soft-X-ray harmonic emission in hollow-core fibers. *Science* **280**, 1412–1415 (1998).
- Anisimov, S. I., Kapeliovich, B. L. & Perel'man, T. L. Electron-emission from surface of metals induced by ultrashort laser pulses. *Sov. Phys. JETP* **39**, 375–377 (1974).
- Cavalleri, A. *et al.* Evidence for a structurally-driven insulator-to-metal transition in VO₂: a view from the ultrafast timescale. *Phys. Rev. B* **70**, 161102(R) (2004).
- Wall, S. *et al.* Quantum interference between charge excitation paths in a solid-state Mott insulator. *Nature Phys.* advance online publication, doi:10.1038/nphys1831 (5 December 2010).
- Holy, J. A., Woo, K. C., Klein, M. V. & Brown, F. C. Raman and infrared studies of superlattice formation in TiSe₂. *Phys. Rev. B* **16**, 3628–3637 (1977).
- El Sayed, K., Schuster, S., Haug, H., Herzel, F. & Henneberger, K. Subpicosecond plasmon response: buildup of screening. *Phys. Rev. B* **49**, 7337–7344 (1994).
- Bányai, L., Vu, Q. T., Mieck, B. & Haug, H. Ultrafast quantum kinetics of time-dependent RPA-screened Coulomb scattering. *Phys. Rev. Lett.* **81**, 882–885 (1998).
- Huber, R. *et al.* How many-particle interactions develop after ultrafast excitation of an electron-hole plasma. *Nature* **414**, 286–289 (2001).
- Sokolowski-Tinten, K. *et al.* Femtosecond X-ray measurement of coherent lattice vibrations near the Lindemann stability limit. *Nature* **422**, 287–289 (2003).
- Siwick, B. J., Dwyer, J. R., Jordan, R. E. & Miller, R. J. D. An atomic-level view of melting using femtosecond electron diffraction. *Science* **302**, 1382–1385 (2003).
- Miaja-Avila, L. *et al.* Laser-assisted photoelectric effect from surfaces. *Phys. Rev. Lett.* **97**, 113604 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements M.B. and S.M. thank M. Aeschlimann for support and discussion. M.B. and S.M. also thank M. Murnane and H. Kapteyn for their support through the NSF EUV ERC. A.C. acknowledges support from the JILA Physics Frontier Center. This work was supported by the German Science Foundation (DFG) within the SFB 855 (C.S., M.B., L.K., K.R.) and by the European Community's FP7 under Marie Curie International Outgoing Fellowship GA 253316 (S.M.). Operation of the Advanced Light Source is supported by the US Department of Energy, Office of Basic Energy Sciences.

Author Contributions M.B. and K.R. conceived the experiment and wrote the paper. T.R., S.H., M.W., B.S., S.M., L.K., M.B. and K.R. realized the experimental time-resolved ARPES setup. A.C., L.M.A. and Y.L. designed and fabricated the EUV multilayer mirrors. T.R., S.H., M.W., C.S. and A.S. collected the time-resolved photoemission data and performed the data analysis. M.K. and K.R. collected and analysed the static photoemission data at the Advanced Light Source. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.B. (bauer@physik.uni-kiel.de; experimental technique) or K.R. (rosnagel@physik.uni-kiel.de; experimental data).

METHODS

Sample preparation and photoemission set-up. 1T-TiSe₂ single crystals were grown from the elements by chemical vapour transport using iodine as transport agent. Samples were mounted on a cryogenic manipulator and cleaved *in situ* under ultrahigh vacuum (UHV) conditions at a base pressure of 3×10^{-10} mbar. Conventional ARPES measurements were conducted at beamline 7.0.1 of the Advanced Light Source at Berkeley, with a Scienta R4000 electron spectrometer. The photon energy was 119 eV and the overall energy resolution was ~ 30 meV. Femtosecond time-resolved ARPES measurements were conducted at the University of Kiel. Photoemitted electrons were detected using a hemispherical electron energy analyzer (SPECS, Phoibos 150) equipped with a two-dimensional detection unit for parallel energy and momentum detection. The total energy resolution of the experiment was mainly governed by the spectral broadening of the femtosecond XUV pulses and was determined to be ~ 400 meV. An independent characterization of the high harmonic radiation with a grating spectrometer showed that the spectral width of the used 27th harmonic was 340 ± 40 meV. Typical integration times for analysis-grade spectra were 3 min. High-quality data as shown in Fig. 2 required an integration time of 15 min.

Pulsed XUV light source. The light source used for the pump–probe photoemission experiments was a 3-kHz Ti:sapphire amplifier system (Dragon, KMLabs, pumped by an Empower 30, Spectra Physics) delivering infrared pulses at 790 nm, 1.2 mJ pulse energy and 32 fs pulse duration. For photoemission, 80% of the pulse energy was used to generate high harmonic femtosecond XUV pulses in an argon-filled hollow-fibre waveguide (XUUS, KMLabs). A pair of multilayer mirrors (total reflectivity, 13%) selected the 27th harmonic ($h\nu = 43$ eV) out of the harmonic spectrum and focused it at an angle of 45° onto the sample mounted in the UHV system. The intensity of the 27th harmonic at the sample position was measured *in situ* using a calibrated XUV photodiode (SXUV 20 HS1, International Radiation Detections) yielding a fluence in the 10^9 photons s^{-1} regime. The residual 20% of the amplifier output was available for infrared photoexcitation of the 1T-TiSe₂ sample. The temporal pulse profile of the pump beam was characterized using the frequency-resolved optical gating technique. The pulse-width of the XUV pulses was estimated to a value < 10 fs from the LAPE infrared–XUV cross-correlation traces of the Se 4p band signal of the 1T-TiSe₂ sample³⁰ (Fig. 2f).

Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3

Zheng Zhou^{1*}, Hanqiao Feng^{1*}, Bing-Rui Zhou¹, Rodolfo Ghirlando², Kaifeng Hu³, Adam Zwolak⁴, Lisa M. Miller Jenkins⁵, Hua Xiao¹, Nico Tjandra⁴, Carl Wu¹ & Yawen Bai¹

The centromere is a unique chromosomal locus that ensures accurate segregation of chromosomes during cell division by directing the assembly of a multiprotein complex, the kinetochore¹. The centromere is marked by a conserved variant of conventional histone H3 termed CenH3 or CENP-A (ref. 2). A conserved motif of CenH3, the CATD, defined by loop 1 and helix 2 of the histone fold, is necessary and sufficient for specifying centromere functions of CenH3 (refs 3, 4). The structural basis of this specification is of particular interest. Yeast Scm3 and human HJURP are conserved non-histone proteins that interact physically with the (CenH3–H4)₂ heterotetramer and are required for the deposition of CenH3 at centromeres *in vivo*^{5–13}. Here we have elucidated the structural basis for recognition of budding yeast (*Saccharomyces cerevisiae*) CenH3 (called Cse4) by Scm3. We solved the structure of the Cse4-binding domain (CBD) of Scm3 in complex with Cse4 and H4 in a single chain model. An α -helix and an irregular loop at the conserved amino terminus and a shorter α -helix at the carboxy terminus of Scm3(CBD) wraps around the Cse4–H4 dimer. Four Cse4-specific residues in the N-terminal region of helix 2 are sufficient for specific recognition by conserved and functionally important residues in the N-terminal helix of Scm3 through formation of a hydrophobic cluster. Scm3(CBD) induces major conformational changes and sterically occludes DNA-binding sites in the structure of Cse4 and H4. These findings have implications for the assembly and architecture of the centromeric nucleosome.

Unlike other eukaryotic species that have complex regional centromeres with multiple centromeric nucleosomes¹⁴, budding yeast has a single centromeric nucleosome that is necessary and sufficient to mediate the accurate segregation of chromosomes during mitosis and meiosis^{15–18}. The simple centromeres of budding yeast provide an attractive system for investigating outstanding topics in centromere biology, including the pathway of CenH3 deposition and the architecture of the centromeric nucleosome^{19,20}.

Yeast Scm3 and human HJURP are binding partners of CenH3–H4 and are functionally required for their deposition at centromeres *in vivo*^{5–13}. A conserved domain of Scm3 dictates specific and stoichiometric binding of CenH3–H4 (Fig. 1a), forming a (Scm3–Cse4–H4)₂ hexamer in 2 M NaCl⁵. This CBD of Scm3 is mapped to residues 84–187 (ref. 5). To investigate the structural basis for the recognition of Cse4 by Scm3, we first analysed the CBD of Scm3 by NMR and found that it is intrinsically disordered (Supplementary Fig. 1). To overcome instability inherent in complexes of individual Scm3, Cse4 and H4 fragments (Supplementary Figs 2–4), we engineered a single-chain molecule in which Scm3 is inserted between Cse4 and H4 to assemble a stably folded molecule (Supplementary Fig. 5). For convenience, we termed the single-chain molecule scSCH (Scm3, Cse4, H4).

The structure of scSCH was determined using multidimensional NMR and verified by structural analysis of its mutants (Fig. 1b–d and Supplementary Fig. 6). The structure of the folded core of scSCH, which includes residues 97–135 of Scm3, 157–202 of Cse4 and 50–99 of H4, is

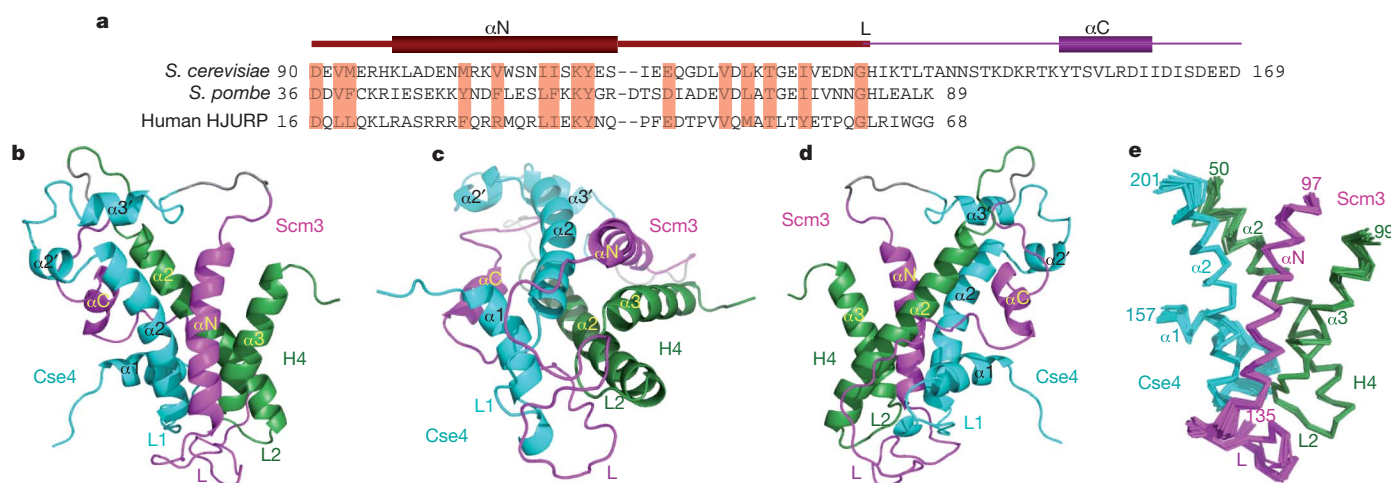


Figure 1 | Overall structure of scSCH. a, The amino acid sequence and secondary structures of the Cse4-binding domain of Scm3 in scSCH. Also shown are the conserved regions in Scm3 of *S. pombe* and human HJURP. Highly conserved residues are highlighted in red. The region in the folded core

is shown in dark magenta (see **e**). **b–d**, Front, bottom and back view of the scSCH structure shown in ribbon representation. Scm3, Cse4 and H4 are in magenta, cyan and dark green, respectively. The full sequence of scSCH is M-His₆-KK-Cse4(150–227)-LVPRGS-Scm3(93–169)-GDK-H4(42–103).

¹Laboratory of Biochemistry and Molecular Biology, National Cancer Institute, Bethesda, Maryland 20892, USA. ²Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892, USA. ³National Magnetic Resonance Facility at Madison, University of Wisconsin, Madison, Wisconsin 53706, USA. ⁴Laboratory of Molecular Biophysics, National Heart, Lung, and Blood Institute, NIH, Bethesda, Maryland 20892, USA. ⁵Laboratory of Cell Biology, National Cancer Institute, NIH, Bethesda, Maryland 20892, USA.

*These authors contributed equally to this work.

well defined with root mean squared deviations (r.m.s.d.) of 0.54 Å for backbone atoms and 1.06 Å for all heavy atoms (Fig. 1e and Supplementary Table 1). Importantly, linker residues inserted between Scm3, Cse4 and H4 do not alter the structure of the folded region. Proteolytic cleavage of the two linkers in the folded scSCH only affects chemical shifts of neighbouring residues close to the cutting sites (Supplementary Fig. 7). Moreover, the folded structure of the above tertiary complex is unchanged by refolding after denaturation in 6 M GdmCl to liberate the three components as individual polypeptides (Supplementary Fig. 8). Backbone amide ^{15}N - $\{^1\text{H}\}$ heteronuclear Overhauser effects (NOE) reflect dynamic motions. The folded core shows small dynamic motions (NOE >0.7) except for loop 1 of Cse4 and the small loop region following the N-terminal α -helix (αN) in Scm3 (Supplementary Fig. 9a, b). In contrast, other regions display larger dynamic motions (NOE <0.7), corresponding to less-well-defined structures (Supplementary Figs 9–11).

In the structure of scSCH, Scm3 interacts broadly with both Cse4 and H4. The αN helix of Scm3 makes close contacts with both the $\alpha 2$ helix of Cse4 and the $\alpha 3$ helix of H4 through multiple hydrophobic interactions (Figs 1b and 2a–d and Supplementary Figs 9c and 13a). Following the αN helix, the loop region of Scm3 (residues 121–144) mainly interacts with loop 1 of Cse4, except that a bulge (Scm3 residues 125–130) in the middle of the loop lies on top of loop 2 of H4 (Fig. 1c and Supplementary Figs 9d and 13b–d). Scm3 loop residues 140–144 also interact with the C-terminal portion of the $\alpha 2$ helix of H4 (Supplementary Figs 9d and 13e). Interestingly, Scm3 residues 145–154 are

completely disordered (Supplementary Figs 9b, e and 10). Finally, the C-terminal α -helix (αC) of Scm3 (155–161) makes interactions with the N-terminal region of the $\alpha 2$ helix of H4 (Supplementary Fig. 13f).

Next, we analysed the effects of mutations on the formation of Scm3–Cse4–H4 complexes with isothermal titration calorimetry. The results reveal that the Scm3 recognition motif resides in the N-terminal region (181–190) of the $\alpha 2$ helix of Cse4. Double mutations Met181Ser/Met184Gly and Ala189Ser/Ser190Val in Cse4 that change the Cse4-specific residues to the corresponding residues in H3 reduced the binding affinity by a factor of 5.5 and 9, respectively (Fig. 2a, b, Supplementary Table 2 and Supplementary Fig. 14). A double mutation Ile110Asp/Ile117Asn in the αN helix of Scm3 decreased the binding affinity by a factor of 85 (Fig. 2a, Supplementary Fig. 14 and Supplementary Table 2). These residues are important for cell growth: mutation of the three residues (Met184, Ala189, and Ser190) in Cse4 to corresponding residues in H3 leads to growth defect (small colony)²¹, and mutation Ile110Asp/Ile117Asn in Scm3 abrogates cell viability⁷, consistent with the effects of these mutations on the binding affinity between Scm3 and Cse4/H4 (Supplementary Table 2). Met 181 should also be important for cell function because it interacts with Ile 117 of Scm3 (Fig. 2a). It is possible that simultaneous mutation of the four residues in Cse4 to the corresponding residues in H3 would abrogate cell viability.

In contrast, deleting the three extra residues Lys 172, Asp 173 and Gln 174 and mutating Thr 170 in loop 1 (to Lys, as in H3), all residues specific to Cse4 (Supplementary Fig. 12), had little effect on the binding affinity (a factor of 1.1) (Supplementary Table 2). Mutation of four residues (Val 165, Thr 166, Asp 167, Glu 168) at the C-terminal region of the $\alpha 1$ helix of Cse4 to corresponding residues in H3 (Ile, Ala, Gln and Asp) also showed little effect on the binding affinity (a factor of 1.4) (Supplementary Table 2). In addition, we found that Scm3 is capable of pulling down the H3^{CATD}–H4 chimera, in which the CATD of Cse4 is swapped to the corresponding region of H3 (ref. 22) (Supplementary Fig. 15). Furthermore, Scm3 can pull down an H3 mutant with only four residues replaced by the corresponding residues in the $\alpha 2$ helix of Cse4 (Met 181, Met 184, Ala 188 and Ser 189) as well (Fig. 2c, d). Importantly, Scm3 residues that interact with the four Cse4-specific residues are well conserved in human HJURP (Fig. 1a). Indeed, Cse4 can also pull down the N-terminal region (residues 2–81) of HJURP (homologous to the Cse4-binding motif of Scm3 (refs 9, 22)) and human H4 (Supplementary Fig. 16). This result is consistent with the ability of Cse4 to replace human CENP-A at centromeres and maintain centromere function in human cells²³. In addition, the CENP-A residues that correspond to the four Cse4-specific residues are reasonably conserved (Fig. 2d).

The structure of scSCH reveals the induction of major local conformational changes in the structure of Cse4 and H4 relative to the (CENP-A–H4)₂ tetramer. First, the packing between the central $\alpha 2$ helices of Cse4 and H4 in the structure of scSCH is loose in comparison to tight hydrophobic interactions in the homology-modelled dimer based on the H3–H4 structure in the nucleosome (Fig. 3a, b), or in the CENP-A–H4 dimer in the (CENP-A–H4)₂ tetramer²⁴. Hydrophobic residues Leu 59, Phe 62 and Val 66 in the $\alpha 2$ helix of H4 lose interacting partners Tyr 193, Ser 192 and Leu 186 in the $\alpha 2$ helix of Cse4 (Fig. 3a, b; Phe 101, Ala 98 and Leu 94 in human CENP-A (ref. 24)). Second, owing to the insertion of the Scm3 loop, loop 1 in Cse4 loses close contact with loop 2 of H4 (Supplementary Fig. 13b) when compared with the corresponding loops in the canonical histone octamer²⁵ (Fig. 3c, d) or in the human (CENP-A–H4)₂ tetramer²⁴ (Supplementary Fig. 17). Third, the $\alpha 2$ helix of Cse4 kinks in the middle in scSCH (Fig. 4a), as forced by the side chain of residue Met 103 of the αN helix of Scm3 (Fig. 4a). In contrast, the $\alpha 2$ helix of CENP-A or H3 is relatively straight in the CENP-A–H4 (ref. 24) (Fig. 4b) or H3–H4 tetramer²⁵ (Supplementary Fig. 18). Fourth, the C-terminal region (94–99) of H4 in scSCH adopts a striking helical conformation and extends the $\alpha 3$ helix of H4. The helical conformation is induced by the side chain

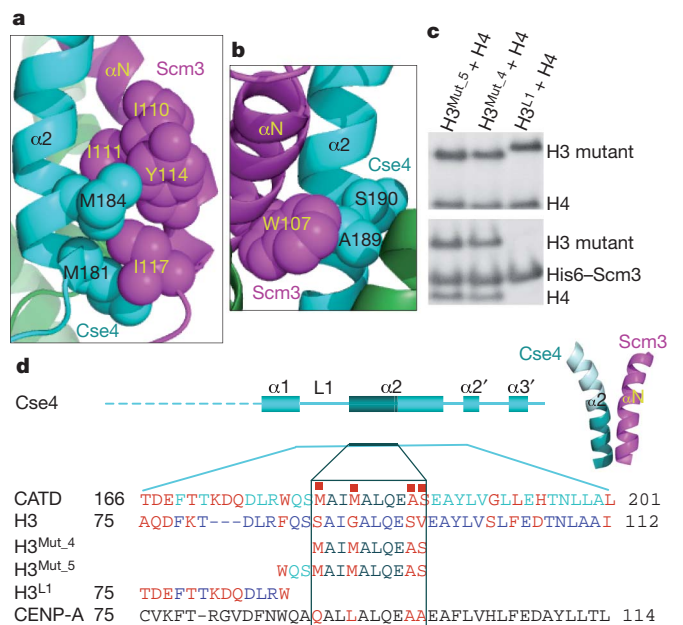


Figure 2 | The N-terminal region (181–190) of the $\alpha 2$ helix of Cse4 is the Scm3 recognition motif. **a**, Ile 110, Tyr 114 and Ile 117 (balls in magenta) in the αN helix of Scm3 form a hydrophobic cluster with Cse4-specific residues Met 181 and Met 184 (balls in cyan) in the $\alpha 2$ helix of Cse4. **b**, Trp 107 (balls in magenta) in the αN helix of Scm3 has close interactions with the Cse4-specific residue Ala 189 (balls in cyan) in the $\alpha 2$ helix of Cse4. Ser 190 is also a Cse4-specific residue (balls in cyan). **c**, SDS-PAGE gels showing the pull-down results with mutants of H3. The top panel shows the input of H3 mutants and H4. The bottom panel shows the molecules eluted from His₆-Scm3 (Scm3(65–169))-bound beads with 250 mM imidazole. H3^{Mut_4}, H3^{Mut_5} and H3^{L1} are the mutants of H3 (see **d**). **d**, Illustration of the secondary structures in Cse4 and the Scm3 recognition motif (dark cyan), CATD, and the mutants used in the pull-down experiments. The red squares indicate the four residues that are sufficient for specific recognition of Cse4 by Scm3. The sequences swapped from Cse4 to H3 in the mutations are shown. The sequences that are not changed in the swap are omitted.

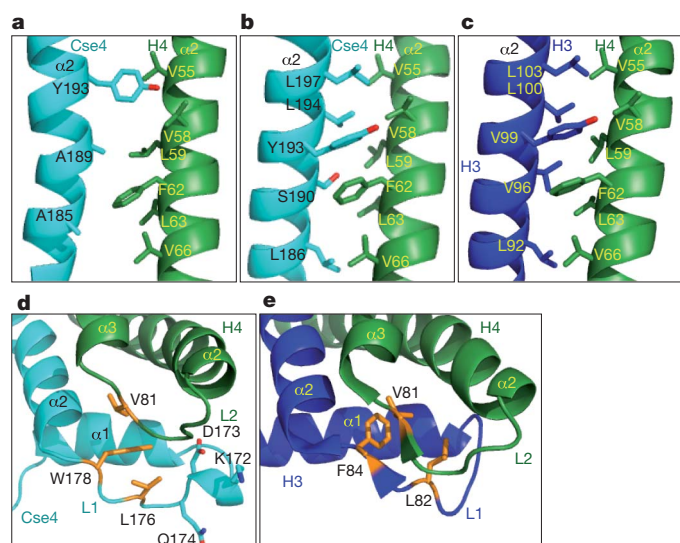


Figure 3 | Altered interactions in the CATD region in scSCH. **a**, The region of the $\alpha 2$ helices of Cse4 and H4 in scSCH, showing that there is little side-chain interaction between the two helices. **b**, The corresponding region of the $\alpha 2$ helices of Cse4 and H4 in the Cse4-H4 dimer structure obtained by homology modelling based on the structures of H3 and H4 in the nucleosome, showing that there are many hydrophobic interactions. **c**, The region of the $\alpha 2$ helices of H3 and H4 in the nucleosome. **d**, Region of loop 1 of Cse4 and the loop 2 of H4. The side chains of the hydrophobic residues are shown in stick representation and orange. The extra three residues in loop 1 of Cse4 are shown in stick representation. **e**, The corresponding loop 1 of H3 and loop 2 of H4 in the nucleosome structure. The hydrophobic residues are shown in stick representation and orange.

of Leu 98 in the αN helix of Scm3 through hydrophobic interactions with the side chains of Leu 98 and Tyr 99 of H4 (Fig. 4a, Supplementary Fig. 14 and Supplementary Table 2). This region is presumably disordered in the (CENP-A-H4)₂ tetramer²⁴ (Fig. 4b). Interestingly, the same region forms a mini β -strand that pairs with a β -strand of H2A in the canonical histone octamer or with a β -strand of histone

chaperone Asf1 in the Asf1-H3-H4 complex^{26,27} (Supplementary Fig. 18).

Furthermore, the C-terminal region of Cse4 has considerable disorder in scSCH (Figs 4a, b). The same region is also disordered in the Cse4-H4 dimer and is not required for Cse4-H4 binding to Scm3 (Supplementary Fig. 5 and Supplementary Fig. 19). This 'tetramerization domain' is well folded in the (CENP-A-H4)₂ tetramer. Structure modelling shows that imposing this folded domain on the corresponding region of Cse4 in scSCH allows association as a (Scm3-Cse4-H4)₂ hexamer without major structural incompatibility (Supplementary Fig. 20), consistent with the existence of (Scm3-Cse4-H4)₂ hexamers in 2 M NaCl (ref. 5). However, in this context, histone topography in the scSCH structure outside the tetramerization domain displays dramatic global conformational changes when compared with the (CENP-A-H4)₂ tetramer, making the modelled (Scm3-Cse4-H4)₂ hexamer incompatible with DNA binding (Supplementary Fig. 20). Moreover, the Scm3 loop in the scSCH structure blocks loop 2 of H4, which makes contacts with DNA in the canonical nucleosome (Fig. 4a, b).

Thus, the structure of scSCH indicates that retention of Scm3 in association with centromere DNA is unlikely to occur via binding of Scm3(CBD) to Cse4/H4, as binding of DNA and Scm3(CBD) to Cse4/H4 is mutually incompatible. Instead, Scm3(CBD) behaves as a specific histone chaperone, and the retention of Scm3 with Cse4/H4 on centromeric DNA requires its distinct DNA-binding domain (H. Xiao and C. Wu, manuscript in preparation).

The structure of scSCH reveals that a subregion within the CATD, including four Cse4-specific residues in the N-terminal region of the $\alpha 2$ helix of Cse4, is necessary and sufficient for specific recognition by Scm3. Thus, the remainder of the CATD of Cse4 should be important for association with other proteins for Cse4 functions. The CBD of Scm3 uses both induced histone conformation changes^{26,27} and direct steric occlusion²⁸ to prevent Cse4-H4 in the Scm3-Cse4-H4 complex from DNA binding (Supplementary Fig. 21). Conversely, Cse4-H4, with a conformation similar to that of CENP-A-H4 in the (CENP-A-H4)₂ tetramer, is unfavourable for Scm3(CBD) binding but favours DNA binding, indicating a competition mechanism for Scm3 and HJURP as CenH3-specific chaperones²².

METHODS SUMMARY

All the proteins used in the present study were overexpressed in *Escherichia coli* and purified using Ni-NTA column (Qiagen), ion exchange, gel filtration and reverse-phase HPLC (Waters). Uniformly isotope-labelled proteins were produced using M9 medium with ¹⁵NH₄Cl, ¹³C-D-glucose and D₂O as the sole source of the isotopes. The molecular weight and stoichiometry of the complex were determined by velocity and equilibrium sedimentation experiments on a Beckman Coulter Proteome XL-I analytical ultracentrifuge at 20.0 °C. The multi-dimensional NMR spectra were collected on Bruker 500, 600, 800 and 900 MHz and Varian 600 and 800 MHz instruments. The structure was calculated using the distance constraints measured by NMR and the program Xplor-NIH. Mutations were made using a quick-change kit. The binding constants were measured on the MicroCal VP-ITC instrument.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 October 2010; accepted 17 January 2011.

Published online 16 March 2011.

- Cleveland, D. W., Mao, Y. & Sullivan, K. F. Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. *Cell* **112**, 407–421 (2003).
- Henikoff, S., Ahmad, K. & Malik, H. S. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**, 1098–1102 (2001).
- Black, B. E. *et al.* Structural determinants for generating centromeric chromatin. *Nature* **430**, 578–582 (2004).
- Black, B. E. *et al.* Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. *Mol. Cell* **25**, 309–322 (2007).
- Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M. M. & Wu, C. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129**, 1153–1164 (2007).

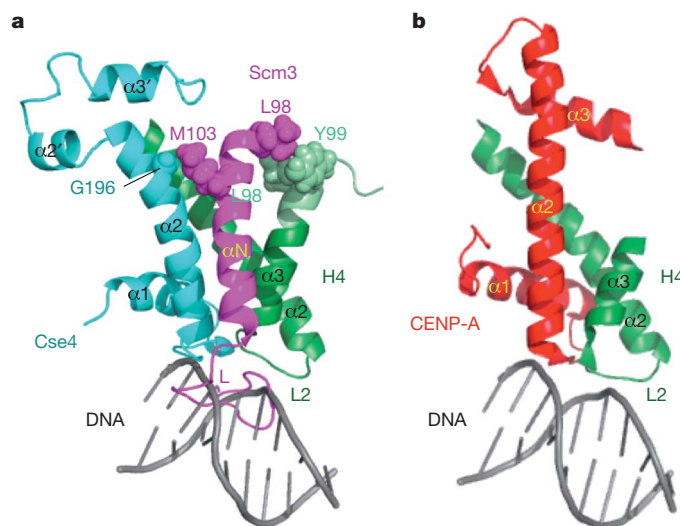


Figure 4 | Scm3 induces large conformational changes in Cse4 and H4 and prevents loop 2 of H4 from binding to DNA. **a**, Cse4-H4 in scSCH. The extended $\alpha 3$ helix in H4 is shown in light green. The loop of Scm3 pushes loop 1 of Cse4 away from loop 2 of H4 and prevents loop 2 of H4 from binding to DNA. DNA is modelled to bind the loop 2 region of H4 based on the canonical nucleosome structure²⁹. **b**, CENP-A-H4 in the (CENP-A-H4)₂ tetramer. DNA is modelled to bind to the loop 2 region of H4 based on the canonical nucleosome structure²⁹.

6. Camahort, R. *et al.* Scm3 is essential to recruit the histone H3 variant Cse4 to centromeres and to maintain a functional kinetochore. *Mol. Cell* **26**, 853–865 (2007).
7. Stoler, S. *et al.* Scm3, an essential *Saccharomyces cerevisiae* centromere protein required for G2/M progression and Cse4 localization. *Proc. Natl Acad. Sci. USA* **104**, 10571–10576 (2007).
8. Camahort, R. *et al.* Cse4 is part of an octameric nucleosome in budding yeast. *Mol. Cell* **35**, 794–805 (2009).
9. Sanchez-Pulido, L., Pidoux, A. L., Pointing, C. P. & Allshire, R. C. Common ancestry of the CENP-A chaperones Scm3 and HJURP. *Cell* **137**, 1173–1174 (2009).
10. Williams, J. S., Hayashi, T., Yanagida, M. & Russell, P. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol. Cell* **33**, 287–298 (2009).
11. Pidoux, A. L. *et al.* Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin. *Mol. Cell* **33**, 299–311 (2009).
12. Foltz, D. R. *et al.* Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. *Cell* **137**, 472–484 (2009).
13. Dunleavy, E. M. *et al.* HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137**, 485–497 (2009).
14. Malik, H. S. & Henikoff, S. Major evolutionary transitions in centromere complexity. *Cell* **138**, 1067–1082 (2009).
15. Stoler, S., Keith, K. C., Curnick, K. E. & Fitzgerald-Hayes, M. A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes Dev.* **9**, 573–586 (1995).
16. Meluh, P. B., Yang, P., Glowczewski, L., Koshland, D. & Smith, M. M. Cse4p is a component of the core centromere of *Saccharomyces cerevisiae*. *Cell* **94**, 607–613 (1998).
17. Cottarel, G., Shero, J. H., Hieter, P. & Hegemann, J. H. A 125-base-pair CEN6 DNA fragment is sufficient for complete meiotic and mitotic centromere functions in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **9**, 3342–3349 (1989).
18. Clarke, L. & Carbon, J. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* **287**, 504–509 (1980).
19. Black, B. E. & Bassett, E. A. The histone variant CENP-A and centromere specification. *Curr. Opin. Cell Biol.* **20**, 91–100 (2008).
20. Furuyama, T. & Henikoff, S. Centromeric nucleosomes induce positive DNA supercoils. *Cell* **138**, 104–113 (2009).
21. Keith, K. C. *et al.* Analysis of primary structural determinants that distinguish the centromere-specific function of histone variant Cse4p from histone H3. *Mol. Cell. Biol.* **19**, 6130–6139 (1999).
22. Shuaib, M., Ouararhni, K., Dimiyrov, S. & Hamiche, A. HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. *Proc. Natl Acad. Sci. USA* **107**, 1349–1354 (2010).
23. Wieland, G., Orthaus, S., Ohndorf, S., Diekmann, S. & Hemmerich, P. Functional complementation of human centromere protein A (CENP-A) by Cse4p from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **24**, 6620–6630 (2004).
24. Sekulic, N., Bassett, E. A., Rogers, D. J. & Black, B. E. The structure of (CENP-A–H4)₂ reveals physical features that mark centromeres. *Nature* **467**, 347–351 (2010).
25. Wood, C. M. *et al.* High-resolution structure of the native histone octamer. *Acta Crystallogr.* **61**, 541–545 (2005).
26. English, C. M., Adkins, M. W., Carson, J. J., Churchill, M. E. & Tyler, J. K. Structural basis for the histone chaperone activity of Asf1. *Cell* **127**, 495–508 (2006).
27. Natsume, R. *et al.* Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature* **446**, 338–341 (2007).
28. Zhou, Z. *et al.* NMR structure of chaperone Chz1 complexed with histones H2A.Z–H2B. *Nature Struct. Mol. Biol.* **15**, 868–869 (2008).
29. White, C. L., Suto, R. K. & Luger, K. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* **20**, 5207–5218 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Ying, K. Varney, J. F. Ellena and J. Gruschus for help collecting NMR spectra, A. Bax for discussion, C. Klee and M. Lichten for comments on the manuscript, and D. Cleveland for plasmids of human CENP-A and H4 histones. This work is supported by the intramural research programs of NCI, NIDDK and NHLBI.

Author Contributions Z.Z. and H.F. contributed equally to this work. Z.Z. performed protein engineering, biochemical and ITC studies. B.-R.Z. contributed to protein sample preparation. B.-R.Z. and L.M.M.J. contributed to the analysis of ITC data. H.F., K.H., A.Z. and N.T. collected the NMR spectra. H.F. and Z.Z. analysed the NMR data and H.F. solved the structure. R.G. performed the sedimentation experiments. H.X. provided initial plasmids and guidance in cloning. C.W. proposed the project and participated in manuscript writing. Y.B. contributed to the overall strategy, project management and writing of the manuscript. All authors read and commented on the manuscript.

Author Information The atomic coordinates have been deposited in the Protein Data Bank under accession code 2L5A. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.B. (yawen@helix.nih.gov).

METHODS

Protein sample preparation. All proteins were expressed in *E. coli* (BL21-codonPlus(DE3)-RIL) with pET vectors (Stratagene). N-terminal His₆-tagged fragments of Cse4 and Scm3 and their mutants were first purified via Ni-NTA (Qiagen) whereas H4 and non-His-tagged fragments of Cse4 and Scm3 were first purified via SP sepharose (GE Healthcare). They were next subjected to reverse-phase HPLC purification using acetonitrile and water as solvents. Purified proteins were lyophilized. Isotope-labelled proteins for NMR studies were produced by growing *E. coli* cells in M9 media with ¹⁵NH₄Cl, U-¹³C₆-glucose, and D₂O as the sole source for nitrogen, carbon and deuterium, respectively. For the measurement of side-chain NOEs, specific methyl labelling (¹³CH₃) for Ile, Leu and Val residues was also made following the protocol of ref. 30.

To prepare the Cse4, H4 and Scm3 complexes, lyophilized proteins were first dissolved in H₂O. Their concentrations were determined by measuring the absorbance at 280 nm. Equal amounts of each species were mixed together and dialysed against 10 mM Tris-HCl and 2 M NaCl at pH 7.4 and 4 °C. After centrifugation, the soluble fractions were subjected to gel filtration on Superdex 75 10/300 GL column (GE Healthcare). The eluted complexes were concentrated with an Amicon with Ultra Ultracel-10 membrane (Millipore) and exchanged to a final buffer of 50 mM MES at pH 5.6. The Cse4-H4 complexes were made in the same way. Scm3(80–211) samples for NMR study were prepared by dissolving them in 8 M urea and dialysis against corresponding buffer. scSCH and all other single-chain molecules derived from scSCH are purified with Ni-NTA (Qiagen) under native conditions (20 mM Tris-HCl and 0.5 M NaCl at pH 8.0), followed by gel filtration with Superdex 200 10/60 column at 4 °C (GE healthcare). The fractions containing the target protein were combined and concentrated and exchanged with the final buffer (50 mM MES at pH 5.4).

Analytical ultracentrifugation. Sedimentation velocity experiments were conducted in duplicate at 20.0 °C on a Beckman Coulter Proteome XL-I analytical ultracentrifuge. 400 µl of the sample of 35 µM in 50 mM MES (pH 5.6) was loaded in a double sector centrepiece cell and analysed at a rotor speed of 50,000 r.p.m. One-hundred scans were acquired as single absorbance measurements ($\lambda = 280$ nm) at 7.1-min intervals using a radial spacing of 0.003 cm. Data were analysed in SEDFIT 11.71 in terms of a continuous $c(s)$ distribution to obtain a sedimentation coefficient, s , and molecular mass M (ref. 31). Solution densities ρ were measured at 20.0 °C on a Mettler Toledo DE51 density meter and solution viscosities η were measured using a Cannon-Ubbelohde viscometer and Cannon-CT 500 constant temperature bath set at 20.00 °C. The partial specific volume v of the complex was calculated in SEDNTERP 1.09 (ref. 32). $c(s)$ analyses were carried out using an s -value range of 0.5 to 6.0 with a linear resolution of 100 and a confidence level (F -ratio) of 0.68. The analyses, implemented using time-independent noise corrections, returned root mean square deviation (r.m.s.d.) values for the best fits of 0.0040 absorbance units. Sedimentation equilibrium experiments were conducted at 20.00 °C on a Beckman Optima XL-A. 135 µl volumes of the complex were studied at loading concentrations of 20, 39 and 78 µM, along with the sample recovered from the sedimentation velocity experiments. Experiments were carried out using six-channel centrepiece cells at rotor speeds ranging from 18,000 to 34,000 r.p.m. In all cases data were acquired as an average of four absorbance measurements at wavelengths of 280 and 250 nm using a radial spacing of 0.001 cm. Sedimentation equilibrium at each speed was achieved within 40 h. Data were analysed globally in terms of a single ideal species using SEDPHAT 6.21 (refs 32, 33).

NMR experiments. NMR experiments were performed on Bruker 500, 600, 800 and 900 MHz and Varian 600 and 800 MHz spectrometers at 35 °C. The following experiments were recorded. 2D: [¹H, ¹H]-NOESY, [¹H, ¹⁵N]-TROSY, [¹H, ¹³C]-HMQC, ¹⁵N-¹H NOE; TROSY version 3D: HNCACB, HNCOCACB, HNCA, HNCOCACB, HNCO, HNCACB; 3D HBHACONH, HCCH-TOCSY, CCH-TOCSY, CCC(CO)NH, [¹H, ¹⁵N]-NOESY-HSQC, [¹H, ¹⁵N]-NOESY-HSQC

([¹³C]methyl-labelled sample), [¹H, ¹³C]-NOESY-HSQC, [¹H, ¹³C]-NOESY-HSQC ([¹³C]methyl-labelled sample). The spectra were processed using NMRPipe³⁴ and analysed with NMRView³⁵.

Structure calculation. Structure calculation was done using the program Xplor-NIH³⁶. The NOE-derived restraints were subdivided into four classes, strong, medium, weak and very weak, by comparison with NOEs of protons separated by known distances as described previously³⁷. Backbone dihedral angle restraints (ϕ and ψ angles) were obtained from analysis of ¹H α , HN, ¹³C α , ¹³C β , ¹³CO and ¹⁵N chemical shifts by using the program TALOS³⁸. Two constraints per hydrogen bond (dNH-O \leq 2.2 Å and dN-O \leq 3.2 Å) were added in the final structure calculation after initial NOE-derived structures were obtained. The program PROCHECK-NMR³⁹ was used to evaluate the quality of the calculated structures.

Isothermal titration calorimetric experiments. The ITC experiments were performed on a MicroCal VP-ITC by injecting Scm3(83–169) solution (250 µM) to a solution of single-chain Cse4-H4 or their mutants (His₆-KK-Cse4(151–207)-LVPRGS-H4(45–103)) (20 µM) in a chamber of 1.4 ml at 25 °C in 50 mM MES (pH 5.4) and 0.1 M NaCl. Twenty-nine injections (each of 10 µl) were made and the heat released was analysed. The data were analysed as described previously⁴⁰.

Pull-down experiments. Pull-down experiments were carried out in 50 mM sodium phosphate, 25 mM imidazole, 2 M NaCl, pH 8.0 at room temperature. Ni-NTA (Qiagen) beads were mixed with His₆-Scm3(66–169) with a final concentration of 6 µM. Approximately 10-fold excess of (Cse4-H4)₂ or (H3-H4)₂ or their mutants was mixed with beads and incubated. The incubation was at 25 °C for 30 min. The beads were washed with the same buffer three times. The complex formed on the beads was eluted with 250 mM imidazole and analysed by SDS-PAGE. Beads without His₆-Scm3 were also incubated with corresponding (Cse4-H4)₂ under identical conditions to assess background binding and the integrity of the tetramer. No nonspecific binding was identified in 2 M NaCl. For molecules derived from thrombin-digested single-chain proteins, the complex was incubated with Ni-NTA (Qiagen) beads at 25 °C for 30 min and then washed three times. The final complex formed on the beads was eluted with either 8 M urea or 250 mM imidazole. The eluted molecules were analysed by SDS-PAGE.

30. Tugarinov, V., Kanelis, V. & Kay, L. E. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature Protocols* **1**, 749–754 (2006).
31. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
32. Cole, J. L., Lary, J. W., Moody, T. P. & Laue, T. M. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol.* **84**, 143–179 (2008).
33. Schuck, P. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal. Biochem.* **320**, 104–124 (2003).
34. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
35. Johnson, B. A. & Blevins, R. A. NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603–614 (1994).
36. Schwieters, C. D., Kuszewski, J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
37. Zhou, Z. *et al.* NMR structure of chaperone Chz1 complexed with histones H2A.Z-H2B. *Nature Struct. Mol. Biol.* **15**, 868–869 (2008).
38. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302 (1999).
39. Laskowski, R. A. *et al.* AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
40. Houtman, J. C. *et al.* Binding specificity of multiprotein signaling complexes is determined by both cooperative interactions and affinity preferences. *Biochemistry* **43**, 4170–4178 (2004).

An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress

Hidetaka Ito^{1†}, Hervé Gaubert^{1*}, Etienne Bucher^{1*†}, Marie Mirouze^{1†}, Isabelle Vaillant^{1†} & Jerzy Paszkowski¹

Eukaryotic genomes consist to a significant extent of retrotransposons that are suppressed by host epigenetic mechanisms, preventing their uncontrolled propagation^{1,2}. However, it is not clear how this is achieved. Here we show that in *Arabidopsis* seedlings subjected to heat stress, a *cop*ia-type retrotransposon named *ONSEN* (Japanese 'hot spring') not only became transcriptionally active but also synthesized extrachromosomal DNA copies. Heat-induced *ONSEN* accumulation was stimulated in mutants impaired in the biogenesis of small interfering RNAs (siRNAs); however, there was no evidence of transposition occurring in vegetative tissues. After stress, both *ONSEN* transcripts and extrachromosomal DNA gradually decayed and were no longer detected after 20–30 days. Surprisingly, a high frequency of new *ONSEN* insertions was observed in the progeny of stressed plants deficient in siRNAs. Insertion patterns revealed that this transgenerational retrotransposition occurred during flower development and before gametogenesis. Therefore in plants with compromised siRNA biogenesis, memory of stress was maintained throughout development, priming *ONSEN* to transpose during differentiation of generative organs. Retrotransposition was not observed in the progeny of wild-type plants subjected to stress or in non-stressed mutant controls, pointing to a crucial role of the siRNA pathway in restricting retrotransposition triggered by environmental stress. Finally, we found that natural and experimentally induced variants in *ONSEN* insertions confer heat responsiveness to nearby genes, and therefore mobility bursts may generate novel, stress-responsive regulatory gene networks.

In *Arabidopsis* mutants compromised in 24-nucleotide siRNA biogenesis, transposon transcripts appear but transposition has not been observed^{3,4}. This is in contrast to mutants lacking chromatin-remodelling factor DDM1 or DNA methyltransferase MET1, in which transposons move during inbreeding^{5–9}. It has been shown that transposon transcripts and their siRNAs accumulate in the vegetative nucleus of pollen¹⁰. A similar observation has been reported for the endosperm^{11–13}. For pollen vegetative cells, where transposon mobility has been observed, it has been postulated that relocation of transposon siRNAs to sperm cells contributes to transposon silencing in the germ line¹⁰. Nevertheless, it is troubling that transposons remain immobile during inbreeding of mutants affected in siRNAs biogenesis, questioning the role of siRNAs in the control of germinal and, therefore, transgenerational transposon mobility¹¹.

We showed previously that a temperature shift applied to 1-week-old seedlings transiently destabilized transcriptional gene silencing (TGS) at loci residing within constitutive heterochromatin where TGS was re-established during the next 24 h¹⁴. A notable exception was a *Ty1/copia*-type long terminal repeat (LTR) retrotransposon family (*ATCOPIA78*), which retained high levels of transcripts two days later¹⁴. This was also observed in older plants (21 days) subjected to raised temperatures¹⁵. In the genome of the Columbia accession, *ATCOPIA78* consists of eight members (Supplementary Fig. 1a),

hereafter referred to as *ONSEN*, of which three have identical LTR sequences, indicating recent transposition (Supplementary Fig. 1b).

By northern blotting, we compared *ONSEN* transcripts in seedlings subjected to a temperature shift of 24 h at 6 °C followed by 24 h at 37 °C (hereafter called heat stress (HS)) to transcripts of seedlings subjected to a control stress (CS) of 24 h at 6 °C followed by 24 h at 21 °C (Fig. 1a). *ONSEN* transcripts were detected in HS plants directly after the stress treatment and for up to 3 days of recovery at 21 °C (HS+3). The longest RNA found corresponded to the full-length transposon (Fig. 1a), whereas smaller RNAs appeared to belong to aberrant RNAs often associated with transcriptionally activated retroelements^{9,16}. Full-length *ONSEN* transcripts were not observed in plants subjected to CS or in non-stressed plants (Fig. 1a and data not shown).

To examine further the specificity of TGS release and to determine possible epigenetic mechanisms involved in *ONSEN* control, we tested RNAs of plants treated with DNA methylation inhibitor 5-azacytidine (AzaC, Fig. 1b) and of *ddm1* mutant plants (Fig. 1c). Neither AzaC treatment nor *ddm1* mutation was effective for transcriptional activation of *ONSEN*, indicating that a reduction of DNA methylation is not sufficient for releasing *ONSEN* silencing. Furthermore, we applied HS and CS treatments to mutants compromised in epigenetic regulation (Fig. 1c, d). We examined *ddm1* mutants (Fig. 1c) and mutants affected in siRNA biogenesis (Fig. 1d): *nrpd1* (ref. 3), impaired in plant-specific RNA polymerase IV (PolIV); *nrpd2* (ref. 17), impaired in the common subunit of RNA PolIV and PolV; *rdr2* (ref. 18), impaired in RNA-dependent RNA polymerase 2; and *dcl3* (ref. 18), mutated in Dicer-like 3. We also challenged the *svh2* (ref. 19) mutant (Fig. 1d), which is deficient in a putative histone 3 lysine 9 methyltransferase. *ONSEN* transcripts were only observed in RNA samples after HS but not after CS treatment (Fig. 1a, c, d). Their levels were not affected by the *ddm1* mutation (Fig. 1c). In contrast, HS-induced accumulation of *ONSEN* RNA was significantly higher in *nrpd1*, *nrpd2*, *rdr2*, *dcl3* and *svh2* mutants (Fig. 1d). During the recovery period following stress, *ONSEN* transcripts diminished and after 10 days (HS+10) the full-length RNAs of the transposon were not detected on northern blots of all genotypes tested (Fig. 1a and Supplementary Fig. 2). These results indicated that siRNA-mediated regulation is responsible for the restriction of *ONSEN* transcript levels after HS, but is not involved in resiliencing during the recovery period.

As PolIV is crucial for the biogenesis of the majority of 24-nucleotide siRNAs⁴, we compared the levels of *ONSEN*-specific siRNAs in wild-type and *nrpd1* plants in relation to the HS-induced accumulation of its transcripts. Noticeably, directly after HS, when *ONSEN* transcript levels were highest, siRNA levels were low and increased only after 1 day of recovery (Fig. 1a, e, f). These siRNAs appeared in both wild-type and *nrpd1* plants and were mainly of the 21-nucleotide siRNA class (Fig. 1e, f), which is thought to direct cleavage of corresponding messenger RNAs. Although levels of 21-nucleotide siRNAs were significantly higher in *nrpd1* mutant plants than in the wild type, massive

¹Department of Plant Biology, University of Geneva, Sciences III, 30 Quai Ernest-Ansermet, CH-1211 Geneva 4, Switzerland. †Present addresses: Division of Biological Sciences, Graduate School of Science, Hokkaido University, 060-0810 Sapporo, Japan (H.I.); Botanical Institute, Hebelstrasse 1, CH-4056 Basel, Switzerland (E.B.); Institut de Recherche pour le Développement (IRD), UMR DIADE, Plant Diversity, Adaptation and Development Laboratory, Université Montpellier 2, 911 Avenue Agropolis, 34394 Montpellier, France (M.M.); Centre National de la Recherche Scientifique (CNRS), UMR 6247 - GReD - INSERM U 931, Clermont Université, 24 avenue des Landais, BP 80026, 63171 Aubière, France (I.V.).

*These authors contributed equally to this work.

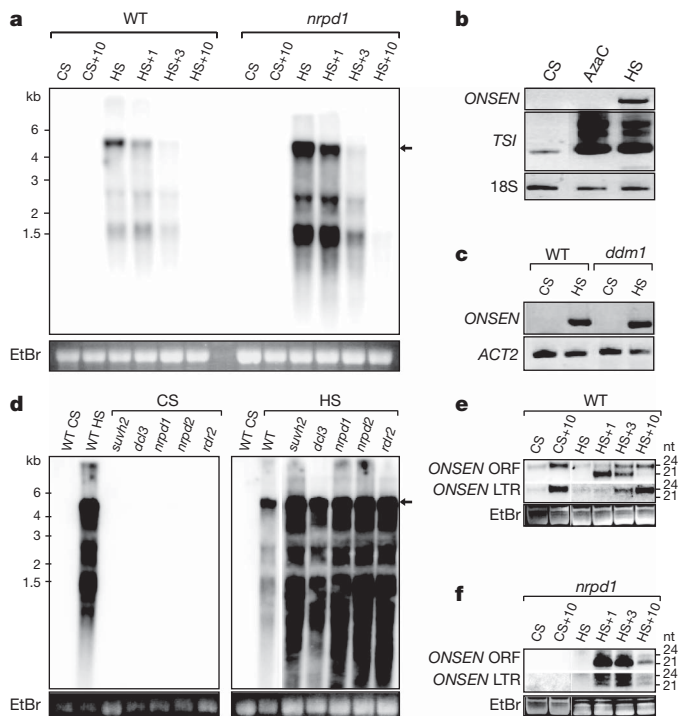


Figure 1 | Heat-stress induced *ONSEN* transcription. **a**, Northern blot revealing *ONSEN* transcripts in wild-type (WT) and *nrpd1* seedlings subjected to HS and after recovering from HS for 1, 3 and 10 days (HS+1, HS+3, HS+10); CS plants were subjected to the control stress (CS). The arrow marks the full-length *ONSEN* transcript. An ethidium bromide (EtBr)-stained gel is shown as loading control. **b**, Detection of *ONSEN* transcripts after AzaC and HS treatments by semiquantitative reverse transcription followed by PCR (RT-PCR) in wild-type plants. *TRANSCRIPTIONALLY SILENT INFORMATION* (TSI) was used as a positive control for the activation of heterochromatic transcription¹⁴; 18S ribosomal RNA was used as an internal control. **c**, Levels of *ONSEN* transcripts in *ddm1* mutant plants subjected to CS or HS and quantified by RT-PCR with *ACTIN2* transcripts (*ACT2*) as an internal control. **d**, Northern blots showing *ONSEN* transcript levels in selected mutants (marked above each lane) subjected to CS or HS. The CS blot was overexposed for possible detection of low levels of *ONSEN* transcripts. An EtBr-stained gel below is shown as a loading control. **e**, **f**, Northern blots of *ONSEN* siRNAs derived from open-reading frame (ORF) or LTR regions accumulating in wild-type (**e**) and *nrpd1* mutant plants (**f**). An EtBr-stained gel is shown as a loading control. nt, nucleotide. See Methods for probe information.

amounts of *ONSEN* transcripts were observed. Thus, this siRNA class was not able to prevent the accumulation of transposon-derived mRNA. However, the high background smear visible on northern blots (Fig. 1a, d) and even more apparent on overexposed blots of RNA isolated from mutants impaired in the siRNA pathway 10 days after HS (Supplementary Fig. 2) may be indicative of *ONSEN* transcript degradation. In the course of HS recovery (at 3 and 10 days), 24-nucleotide siRNAs highly accumulated in wild-type but not in *nrpd1* plants, and were especially abundant for the LTR regions of *ONSEN* (Fig. 1e, f). However, the contribution of 24-nucleotide siRNAs to resiliencing at *ONSEN* loci during recovery is not clear, as they accumulate also in CS plants grown for an additional 10 days after CS treatment (Fig. 1e). Therefore their levels seem not to be related to the HS treatment. Moreover, *ONSEN* resiliencing also occurred in *nrpd1*, where they were mostly absent (Fig. 1f).

Detection of *ONSEN* full-length transcripts, potentially able to serve as templates for reverse transcription, prompted us to examine the DNA of *Arabidopsis* subjected to HS. By Southern blot analysis we detected a significant increase in *ONSEN* copy number and observed a banding pattern indicative of the presence of two forms of extrachromosomal transposon copies, one linear reflecting the 2.8-kb fragment

and one circular containing a single LTR consistent with the 4.5-kb fragment (Fig. 2a, left). Linear extrachromosomal forms are capable of chromosomal integration, in contrast to the circular forms that have been considered as by-products of retroelement replication^{20,21}. In *nrpd1* and other mutant plants affected in the siRNA pathway, the abundance of HS-induced *ONSEN* DNA was significantly higher than in wild type (Fig. 2a, left). Noticeably, similarly high levels were observed in HS-treated *suwh2* mutant plants (Fig. 2a, left), which are known to exhibit wild-type levels of siRNAs²². After 10 days of recovery, *ONSEN* extrachromosomal DNA was still at a relatively high level but almost exclusively in the linear form (Fig. 2a, right).

Real-time quantitative polymerase chain reaction (qPCR) during HS and after subsequent recovery was performed to determine the kinetics of *ONSEN* DNA accumulation in wild-type plants and in a representative siRNA-biogenesis mutant (*nrpd1*) and to examine possible changes in copy number due to chromosomal integration events (Fig. 2b). Over the first 4 h of the temperature shift from 6 °C to 37 °C, the abundance of *ONSEN* DNA did not change. However, after 6 h, *ONSEN* copy number increased significantly from 8 endogenous copies to more than 30 in *nrpd1* but not in the wild type (Fig. 2b). After 12 h of HS, *ONSEN* copy number had increased in the wild type to more than 25 and in the *nrpd1* mutant to more than 160. The maximal copy numbers of over 50 for the wild type and over 500 for *nrpd1* mutants were reached 12 h and 24 h after HS, respectively. The HS-induced increase in *ONSEN* DNA seemed to be biphasic and this was especially pronounced in *nrpd1* plants (Fig. 2b). So far, we have no explanation for this biphasic accumulation but it may be related to stress-triggered synchronization of the retroelement replication cycle.

During 20–30 days of subsequent growth of both wild-type and *nrpd1* plants *ONSEN* copy number gradually decreased, reaching the initial number of the Columbia accession (Fig. 2b), consistent with the absence of or only sporadic chromosomal integration events. To examine whether new somatic integrations occurred, we performed transposon display on plants grown for 40 days after HS (Supplementary Fig. 3a). New *ONSEN* insertions were not detected in the genomic DNA of either wild-type or *nrpd1* mutant plants, consistent with the qPCR results. However, we can not exclude the possibility of rare transposition events occurring late in vegetative development leading to only small sectors with new insertions.

It has been suggested that transgenerational transposon mobility is suppressed during gametophyte formation by siRNAs^{10,23}. However, there is no evidence based on germinal transposition events to support this hypothesis. By transposon display and Southern blot hybridization,

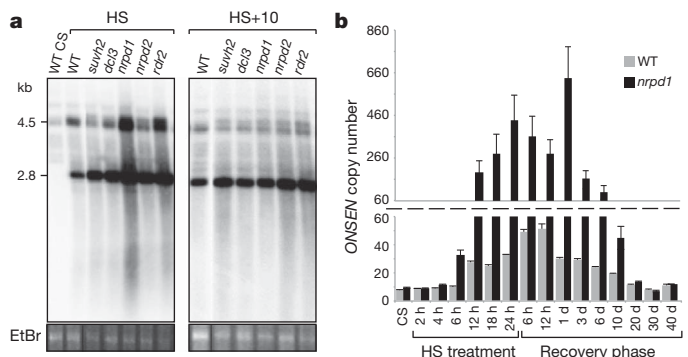


Figure 2 | Accumulation of *ONSEN* extrachromosomal DNA. **a**, Southern blot of Ψ SI-digested DNA isolated from HS-treated seedlings of wild type and selected mutants, directly after HS (left) or after 10 days of recovery (HS+10, right) and hybridized with an *ONSEN*-specific probe (see Methods). A 2.8-kb band is expected for the extrachromosomal linear form of *ONSEN* (Supplementary Fig. 1b). **b**, The kinetics of *ONSEN* DNA accumulation for wild type (grey) and *nrpd1* (black) measured by qPCR during and after HS treatment (mean \pm s.e.m., $n = 3$ biological repetitions).

we examined genomic DNA from the progeny of self-fertilized wild-type and *nprp1* plants subjected to HS and CS for new *ONSEN* insertions. Transposon movement was not detected in the offspring of either *nprp1* or wild-type plants subjected to CS, or in HS-treated wild type (Fig. 3a and Supplementary Fig. 3b). However, a surprisingly high frequency of retrotransposition was recorded in the progeny of *nprp1* mutant plants subjected to HS at the seedling stage (Fig. 3a and Supplementary Fig. 3b, c). Furthermore, the patterns of new *ONSEN* insertions in sibling seedlings derived from a single plant were found to differ in each individual examined, indicating that transposition occurred either before gametogenesis, during gametogenesis, after fertilization, or any combination therein (Fig. 3a and Supplementary Fig. 3b). To distinguish between these alternatives, we analysed *nprp1* progeny plants derived from seeds of different flowers of the same progenitor (Fig. 3b). We found that patterns of new insertions differed entirely between progeny derived from different flowers. However, within the same flower we found common transposition patterns

indicating somatic movement of *ONSEN* during flower development. Moreover, we were not able to find any new and unique *ONSEN* insertions specific to a single plant (Fig. 3b). Therefore all transposition events revealed in the sixteen progeny plants derived from two different flowers must have occurred before the differentiation of male and female gametophytes. Therefore, siRNA-mediated control of retrotransposon movement is not restricted to the gametophytic phase as it has been postulated^{10–13}.

To define better the roles of sporophytic and gametophytic 24-nucleotide siRNAs in suppressing stress-induced transgenerational retrotransposition, we examined heterozygote *nprp1* plants subjected to HS treatment. In these plants the biogenesis of siRNAs is unaffected in the sporophyte, but is deficient in 50% of the male and female gametophytes. We compared the progeny of homozygous *nprp1* mutant plants subjected to HS with homozygous *nprp1* mutant

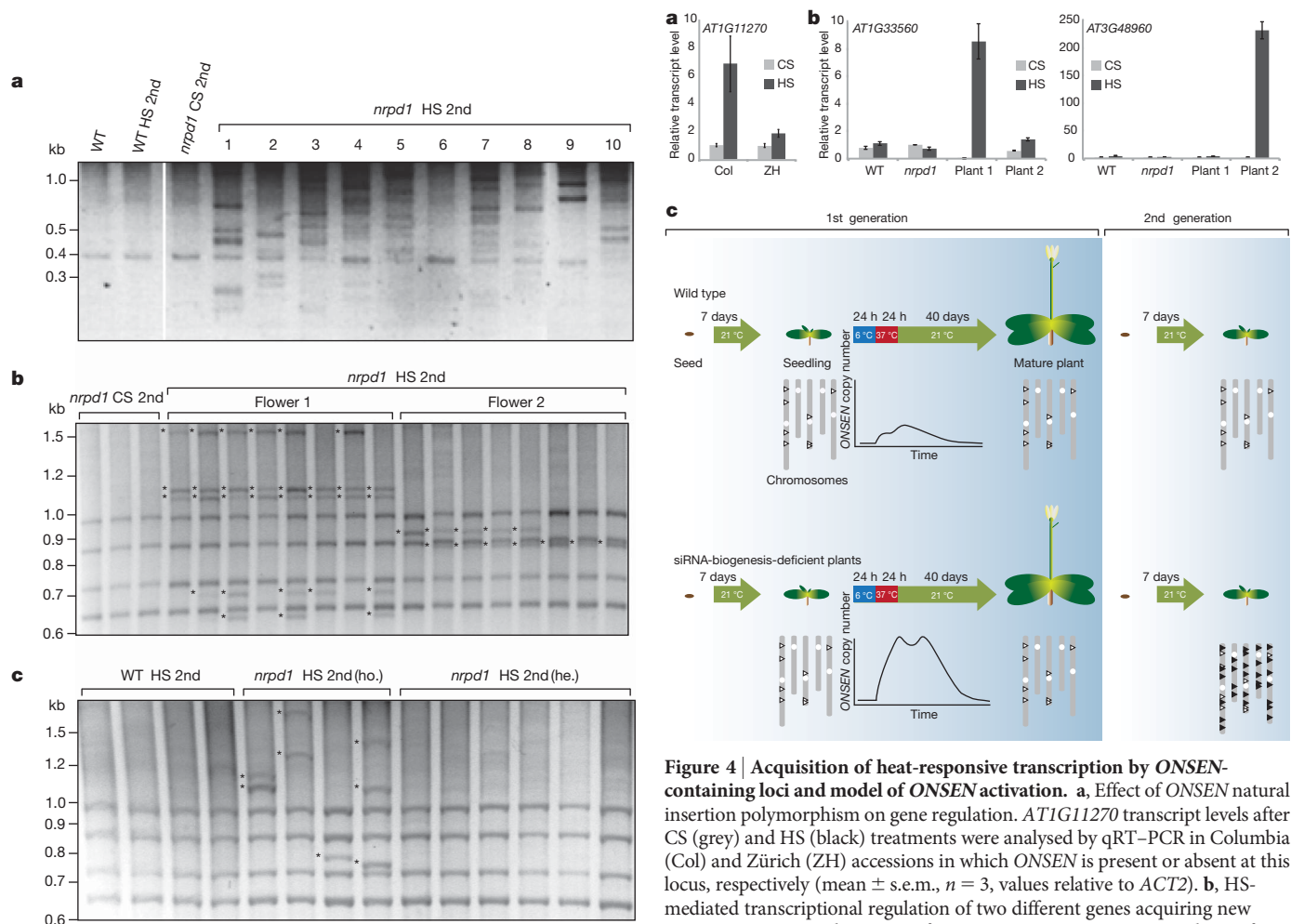


Figure 3 | Burst of *ONSEN* transposition in the progeny of HS-treated *nprp1* plants. **a**, Transposon display (using primer Copia78 3' LTR, Supplementary Table 2) detecting new *ONSEN* insertions. Numbers above the lanes of *nprp1* HS 2nd (second generation) represent 10 individual plants that are siblings derived from bulk-harvested seeds of one *nprp1* plant that was HS-treated as a 7-day-old seedling. **b**, Transposon display (using primer *ONS_312_R*, Supplementary Table 2) detecting new *ONSEN* insertions. Sixteen *nprp1* HS 2nd plants are derived from two flowers of a single HS-treated *nprp1* plant (flower 1 and flower 2 represented by eight plants each). Asterisks mark new *ONSEN* insertions. **c**, Transposon display (with primer as in **b**) revealing new *ONSEN* insertions in the progeny of HS-treated *nprp1* homozygote mutant plants (*nprp1* HS 2nd (ho.)) but not in the progeny of wild-type HS-treated plants (WT HS 2nd) or in *nprp1* homozygote mutant progeny of HS-treated *nprp1* heterozygote mutants (*nprp1* HS 2nd (he.)).

Figure 4 | Acquisition of heat-responsive transcription by *ONSEN*-containing loci and model of *ONSEN* activation. **a**, Effect of *ONSEN* natural insertion polymorphism on gene regulation. *AT1G11270* transcript levels after CS (grey) and HS (black) treatments were analysed by qRT-PCR in Columbia (Col) and Zürich (ZH) accessions in which *ONSEN* is present or absent at this locus, respectively (mean \pm s.e.m., $n = 3$, values relative to *ACT2*). **b**, HS-mediated transcriptional regulation of two different genes acquiring new *ONSEN* insertion in the course of our experiments. Two progeny plants of HS-treated *nprp1* homozygote mutant plant, named plant 1 and plant 2, were selected for displaying new homozygous *ONSEN* insertion at two distinct loci, *AT1G33560* and *AT3G48960*, respectively. Acquired transcriptional responses to HS of the affected genes were revealed by qRT-PCR (legend and values as in **a**). **c**, Summary of experimental results illustrating the role of the siRNA pathway in transgenerational control of *ONSEN* mobility. Upper part of the figure represents wild-type control of *ONSEN* activity and lower part illustrates uncontrolled accumulation of *ONSEN* copy number in siRNA-biogenesis-deficient plants. The graphs under the arrows illustrate the kinetics of *ONSEN* DNA accumulation on HS treatment. The open triangles on five *Arabidopsis* chromosomes represent eight endogenous *ONSEN* copies in the Columbia accession. The black triangles illustrate new *ONSEN* insertions found in the second generation. White circles on the chromosomes specify the location of the centromeres.

segregants derived from HS-treated *nprp1* heterozygotes. A high frequency of retrotransposition was only observed in progeny of homozygous *nprp1* mutant plants (Fig. 3c). These results are consistent with the involvement of 24-nucleotide siRNAs in either erasing 'stress memory' during somatic growth and/or suppressing retrotransposition in flower tissues, rather than with epigenetic control of retrotransposon movement during gametogenesis.

To define better the molecular mechanism controlling *ONSEN* transposition primed by HS, we analysed progenies of further HS-treated plants compromised in epigenetic regulation. Because DDM1 and KRYPTONITE (KYP, histone H3 lysine 9 methyltransferase) were previously implicated in transposition control of a related family of retrotransposons^{6,9}, we subjected both mutants to HS and examined the progeny by transposon display. Retrotransposition was not observed (Supplementary Fig. 4). Despite detecting high levels of transposon transcripts in *suvh2* mutants after HS (Fig. 1d) and a significant increase in *ONSEN* copy number (Fig. 2a), no retrotransposition events were found in the next generation (Supplementary Fig. 4). We next investigated mutants deficient in siRNA biogenesis (*nprp2*, *rdr2* and *dcl3*). Transposition events were observed in *nprp2* and *rdr2* (Supplementary Fig. 4), further indicating that biogenesis of siRNAs is crucial for preventing transgenerational mobility of *ONSEN*. As *dcl3* is essential for 24-nucleotide siRNA biogenesis, we predicted that there would be new *ONSEN* insertions in the progeny of HS-treated *dcl3* plants; in fact, no new insertions were detected (Supplementary Fig. 4). Therefore, although DCL3 clearly restricts the levels of *ONSEN* transcripts after HS, it is dispensable for the control of transgenerational transposition. This hints at two steps in *ONSEN* control: restraining levels of its transient transcription/reverse transcription and suppression of transgenerational transposition. As only the first requires DCL3 and SUVH2, the two control steps seem to be, at least in part, mechanistically independent. Given the functional redundancy of dicer-like (DCL) proteins in *Arabidopsis*²⁴, DCL3 may be substituted possibly by another DCL protein(s) at the second control step. It is also possible that transgenerational control of retrotransposition can occur without the involvement of dicer-like activities, as has been described in animals²⁵.

To determine whether *ONSEN* has preferential insertion targets, we characterized 11 new insertion sites and concluded that, although the retroelement inserted genome wide (Supplementary Fig. 5), it showed a clear preference for transcribed gene regions (all 11 insertions), with a further preference for exons (10 insertions) (Supplementary Table 1). Moreover, 2 of 11 insertions were homozygous (data not shown), which is consistent with retrotransposon movement during flower development but before the differentiation of anthers and carpels.

It has been postulated that a burst of transposition helped to shape plant genomes^{26,27} and to modify their transcriptional responses²⁸. Interestingly, a gene in the Columbia accession harbouring a natural insertion of *ONSEN* was identified as being heat responsive²⁹. To determine the physiological relevance of this observation we analysed heat responsiveness of this gene in the Zürich accession where *ONSEN* is absent at this location (Fig. 4a). Indeed, HS-induced transcriptional activation in the Columbia accession was much more pronounced than in the Zürich accession (Fig. 4a). We determined also whether our experimentally induced retrotransposition events, in the second generation of *nprp1* HS-treated plants, had an impact on the transcriptional regulation of endogenous loci harbouring new *ONSEN* insertions. We examined the heat-stress response of two such genes and showed that they became heat responsive when compared to wild-type or *nprp1* first-generation plants (Fig. 4b). Therefore, it can be predicted that after our experimental burst of *ONSEN* transposition different subsets of genes in various progeny plants will acquire such regulatory properties. Now, having established an environmentally inducible system of transgenerational retrotransposition and having revealed the molecular and developmental mechanisms of its control (Fig. 4c), we are in a position to reproduce retrotransposon bursts in a controlled fashion and to determine their adaptive and/or damaging power.

METHODS SUMMARY

Plant material. All mutants used in this study (*dcl3-1* (ref. 18), *ddm1-2* (ref. 30), *nprp1a-3* (ref. 3), *nprp2a-2/2b-1* (ref. 17), *rdr2-2* SALK_059661 (ref. 18), *suvh2* SALK_079574 (ref. 19)) are in the Columbia (Col-0) background.

Stress treatment. Plants were grown in ½ MS medium (0.8% agar, 1% sucrose) in a Percival CU-22L chamber at 21 °C with 12 h light (140 µmol m⁻² s⁻¹) and 12 h dark. After CS or HS treatment (see text), plants were grown at 21 °C in long-day conditions (16 h light). To analyse the progeny of CS- or HS-treated seedlings, plants were transplanted 10 days after CS/HS to soil and grown under long-day conditions.

RNA, DNA, and transposon display analyses. See Methods for details.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 July 2010; accepted 13 January 2011.

Published online 13 March 2011.

- Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nature Rev. Genet.* **8**, 272–285 (2007).
- Lisch, D. Epigenetic regulation of transposable elements in plants. *Annu. Rev. Plant Biol.* **60**, 43–66 (2008).
- Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120 (2005).
- Mosher, R. A., Schwach, F., Studholme, D. & Baulcombe, D. C. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc. Natl Acad. Sci. USA* **105**, 3145–3150 (2008).
- Miura, A. *et al.* Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. *Nature* **411**, 212–214 (2001).
- Tsukahara, S. *et al.* Bursts of retrotransposition reproduced in *Arabidopsis*. *Nature* **461**, 423–426 (2009).
- Johannes, F. *et al.* Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet.* **5**, e1000530 (2009).
- Reinders, J. *et al.* Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes. *Genes Dev.* **23**, 939–950 (2009).
- Mirouze, M. *et al.* Selective epigenetic control of retrotransposition in *Arabidopsis*. *Nature* **461**, 427–430 (2009).
- Slotkin, R. K. *et al.* Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**, 461–472 (2009).
- Mosher, R. A. *et al.* Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature* **460**, 283–286 (2009).
- Hsieh, T.-F. *et al.* Genome-wide demethylation of *Arabidopsis* endosperm. *Science* **324**, 1451–1454 (2009).
- Gehring, M., Bubb, K. L. & Henikoff, S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* **324**, 1447–1451 (2009).
- Tittel-Elmer, M. *et al.* Stress-induced activation of heterochromatic transcription. *PLoS Genet.* **6**, e1001175 (2010).
- Pecinka, A. *et al.* Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. *Plant Cell* **22**, 3118–3129 (2010).
- Hirochika, H., Okamoto, H. & Kakutani, T. Silencing of retrotransposons in *Arabidopsis* and reactivation by the *ddm1* mutation. *Plant Cell* **12**, 357–369 (2000).
- Onodera, Y. *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622 (2005).
- Xie, Z. *et al.* Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2**, e104 (2004).
- Naumann, K. *et al.* Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J.* **24**, 1418–1429 (2005).
- Feuerbach, F., Drouaud, J. & Lucas, H. Retrovirus-like end processing of the tobacco *Tnt1* retrotransposon linear intermediates of replication. *J. Virol.* **71**, 4005–4015 (1997).
- Hirochika, H. & Otsuki, H. Extrachromosomal circular forms of the tobacco retrotransposon *Tto1*. *Gene* **165**, 229–232 (1995).
- Johnson, L. M., Law, J. A., Khattar, A., Henderson, I. R. & Jacobsen, S. E. SRA-domain proteins required for DRM2-mediated *de novo* DNA methylation. *PLoS Genet.* **4**, e1000280 (2008).
- Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Rev. Genet.* **11**, 204–220 (2010).
- Vaucheret, H. Plant ARGONAUTES. *Trends Plant Sci.* **13**, 350–358 (2008).
- Vagin, V. V. *et al.* A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**, 320–324 (2006).
- Piegu, B. *et al.* Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Res.* **16**, 1262–1269 (2006).
- Ammiraju, J. S. S. *et al.* Evolutionary dynamics of an ancient retrotransposon family provides insights into evolution of genome size in the genus *Oryza*. *Plant J.* **52**, 342–351 (2007).
- Naito, K. *et al.* Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* **461**, 1130–1134 (2009).
- Lim, C. J. *et al.* Gene expression profiles during heat acclimation in *Arabidopsis thaliana* suspension-culture cells. *J. Plant Res.* **119**, 373–383 (2006).
- Vongs, A., Kakutani, T., Martienssen, R. A. & Richards, E. J. *Arabidopsis thaliana* DNA methylation mutants. *Science* **260**, 1926–1928 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank all members of the J.P. laboratory and R. L. Fisher for discussions, L. Broger, M. Freyre, J. Nicolet, C. Mégies for technical assistance and P. King and J. Reinders for editing of the manuscript. This work was supported by grants from the Swiss National Science Foundation (31003A-125005) and the European Commission through the AENEAS collaborative project (FP7 226477), the RECBREED consortium (FP7 227190) and The Epigenome Network of Excellence (FP6 503433).

Author Contributions H.I., E.B., H.G., M.M. and J.P. conceived the study. H.I., E.B., H.G., M.M. and I.V. performed the experiments. J.P. wrote the paper with contributions from E.B. and M.M.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.P. (jerzy.paszkowski@unige.ch).

METHODS

Plant material. All mutants used in this study (*dcl3-1* (ref. 18), *ddm1-2* (ref. 30), *npr1a-3* (ref. 3), *npr2a-2/2b-1* (ref. 17), *rdr2-2* SALK_059661 (ref. 18), *suvh2* SALK_079574 (ref. 19)) are in the Columbia (Col-0) background.

Stress treatment. Plants were grown in $\frac{1}{2}$ MS medium (0.8% agar, 1% sucrose) in a Percival CU-22L chamber at 21 °C with 12 h light ($140 \mu\text{mol m}^{-2} \text{s}^{-1}$) and 12 h dark. After CS or HS treatment (see text), plants were grown at 21 °C in long-day conditions (16 h light). To analyse the progeny of CS- or HS-treated seedlings, plants were transplanted 10 days after CS/HS to soil and grown under long-day conditions.

RNA analysis. RNA was isolated from aerial parts of around 20 plants and northern blots, RT-PCR and siRNA analyses were carried out as described previously⁹. Full-length transcripts were detected with *ONSEN*-specific probe A, and siRNAs were detected with *ONSEN*-specific probe B (LTR region) or probe C (ORF). qRT-PCR analyses were performed using the Quantifast Multiplex PCR Kit (Qiagen). RNA levels were determined using TaqMan assays (qPCR thermocycler 7900HT, Applied Biosystems) and normalized using *ACTIN2*. PCR conditions were 95 °C for 5 min followed by 45 cycles alternating 45 s at 95 °C and 45 s at 60 °C. Probe localization and primer details are given in Supplementary Fig. 1b and Supplementary Table 2, respectively.

DNA analysis. Aerial parts from around 20 plants were collected and DNA was isolated with MiniPrep Kit (Qiagen) following the manufacturer's recommendations.

Southern blots were performed as described previously⁹ using *ONSEN*-specific probe (probe C, see Supplementary Fig. 1b and Supplementary Table 2). For qPCR analysis of *ONSEN* DNA copies, the Quantifast Multiplex PCR Kit (Qiagen) was used and *ACTIN2* was used to normalize DNA levels. DNA copy number was determined using TaqMan assays performed in the qPCR thermocycler 7900HT (Applied Biosystems) in a final volume of 10 μl . PCR conditions were 95 °C for 5 min followed by 45 cycles alternating 45 s at 95 °C and 45 s at 60 °C (primer details in Supplementary Table 2).

Transposon display. A simplified transposon display method based on the GenomeWalker Universal kit (Clontech Laboratories) was developed for library construction, with the following modifications. Genomic DNA (300 ng) was digested overnight with the blunt cutting *DraI* restriction enzyme (Promega) in a final volume of 50 μl , using a tenfold enzyme excess compared with the manufacturer's recommendations. After digestion, DNA fragments were purified on a PCR purification column (Qiagen) following the manufacturer's instructions and eluted into 20 μl ; 5 μl was used for overnight ligation at 16 °C in 16 μl with GenomeWalker adaptors. After ligation, DNA was diluted 20-fold and 1 μl used as a template for the PCR reaction. PCR were performed using a primer specific for *ONSEN* (Copia78 3' LTR or ONS_312_R) and a primer specific for the adaptor (GenWalk_AP1). PCR conditions were 5 min at 95 °C, followed by 33 cycles of 30 s at 94 °C, 30 s at 58 °C, 1 min at 72 °C; and a final elongation step of 7 min at 72 °C. PCR products were separated on 3% agarose gels. Primer details are given in Supplementary Table 2.

Controlling inelastic light scattering quantum pathways in graphene

Chi-Fan Chen¹, Cheol-Hwan Park¹, Bryan W. Boudouris^{2,3}, Jason Horng¹, Baisong Geng¹, Caglar Girit¹, Alex Zettl^{1,3}, Michael F. Crommie^{1,3}, Rachel A. Segalman^{2,3}, Steven G. Louie^{1,3} & Feng Wang^{1,3}

Inelastic light scattering spectroscopy has, since its first discovery^{1,2}, been an indispensable tool in physical science for probing elementary excitations, such as phonons³, magnons⁴ and plasmons⁵ in both bulk and nanoscale materials. In the quantum mechanical picture of inelastic light scattering, incident photons first excite a set of intermediate electronic states, which then generate crystal elementary excitations and radiate energy-shifted photons⁶. The intermediate electronic excitations therefore have a crucial role as quantum pathways in inelastic light scattering, and this is exemplified by resonant Raman scattering⁶ and Raman interference^{7,8}. The ability to control these excitation pathways can open up new opportunities to probe, manipulate and utilize inelastic light scattering. Here we achieve excitation pathway control in graphene with electrostatic doping. Our study reveals quantum interference between different Raman pathways in graphene: when some of the pathways are blocked, the one-phonon Raman intensity does not diminish, as commonly expected, but increases dramatically. This discovery sheds new light on the understanding of resonance Raman scattering in graphene. In addition, we demonstrate hot-electron luminescence⁹ in graphene as the Fermi energy approaches half the laser excitation energy. This hot luminescence, which is another form of inelastic light scattering, results from excited-state relaxation channels that become available only in heavily doped graphene.

Graphene, a two-dimensional carbon sheet^{10,11}, is an attractive system in which to explore novel inelastic light scattering phenomena. Raman scattering from few-layer graphene is readily observable, and has been widely used to distinguish layer thickness¹², characterize quality^{13–15}, and probe electron–phonon interactions^{16,17}. At the same time, graphene has unique optical transitions that can be tuned through electrostatic gating^{18–20}. However, using electrostatic doping to control quantum pathways of inelastic light scattering has presented a challenge, because it requires the modification of intermediate excited states with transition energies close to the laser photon energy. Here we achieve this goal by combining highly efficient ion-gel gating (Methods) and near-infrared laser excitation. This combination enables us to control intermediate optical transition pathways and reveal interesting new optical phenomena. We show that, counter-intuitively, the one-phonon Raman signal increases dramatically when certain quantum pathways are blocked. This represents a notable manifestation of quantum interference between different Raman scattering pathways. On the other hand, the two-phonon Raman signal decreases monotonically with reduced quantum channels. In addition, we observe the emergence of hot luminescence⁹ when the Fermi energy (E_F) approaches half the value of the photon excitation energy. This hot luminescence in graphene arises from new dynamic relaxation channels of the photo-excited states.

The excitation pathways in graphene samples are controlled through electrostatic doping using a high-capacitance ion-gel gate dielectric^{21,22}. Figure 1a displays a diagram of a typical device. The carrier concentration in graphene is controlled by the top gate voltage

(V_g). The doping dependence of electrical transport, optical transmission and inelastic light scattering are measured on the same graphene devices. Figure 1b shows the electrical resistance curve of a graphene device, which has a charge neutral point (CNP) at 1.2 V. The resistance decreases from the CNP value on both electron and hole doping. A lower resistance, corresponding to a higher carrier concentration and a larger shift of Fermi energy, is achieved in the hole-doping region; such hole doping in graphene will be the focus of our study.

To determine the gate-induced Fermi energy shift in graphene samples, we use infrared transmission spectroscopy^{18,19}. This approach is based on the fact that an optical transition can be blocked by hole doping when the initial state has an energy higher than the Fermi level and is not occupied (Fig. 1c). Figure 1d displays the gate-induced change in the real part of graphene optical conductivity ($\Delta\sigma'$) at different gate biases. The spectra exhibit distinct step-like behaviour,

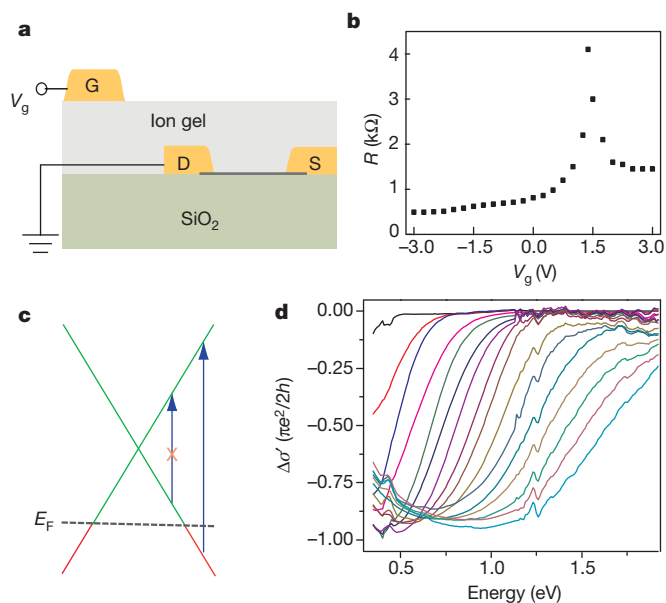


Figure 1 | Controlling the optical transitions in graphene with ion-gel gating. **a**, Illustration of an ion-gel-gated graphene device on a SiO_2 substrate. G, gate; D, drain; S, source. **b**, Graphene resistance (R) as a function of gate voltage, V_g ; R has a maximum at the charge neutral point $V_{\text{CNP}} = 1.2$ V. The largest resistance decrease, corresponding to the highest carrier density, occurs at $V_g = -3$ V with hole doping. **c**, Illustration of gate-induced change in graphene absorption. Electronic states are filled (red line) below E_F and empty (green line) above E_F . Electronic transitions (blue arrows) from zero energy to $2|E_F|$ get blocked by hole doping due to empty initial states. **d**, Change in the real part of optical conductivity, $\Delta\sigma'$, when $V_g - V_{\text{CNP}}$ is varied from -0.25 to -4 V in steps of 0.25 V. The optical conductivity is reduced for optical excitation energies below $2|E_F|$, which increases with gating strength. $2|E_F|$ values up to 1.8 eV are achieved using the ion-gel gate. (The feature around 1.3 eV is an artefact due to a minimum spectral density of our lamp light source at this energy.)

¹Department of Physics, University of California at Berkeley, Berkeley, California 94720, USA. ²Department of Chemical and Biomolecular Engineering, University of California at Berkeley, Berkeley, California 94720, USA. ³Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

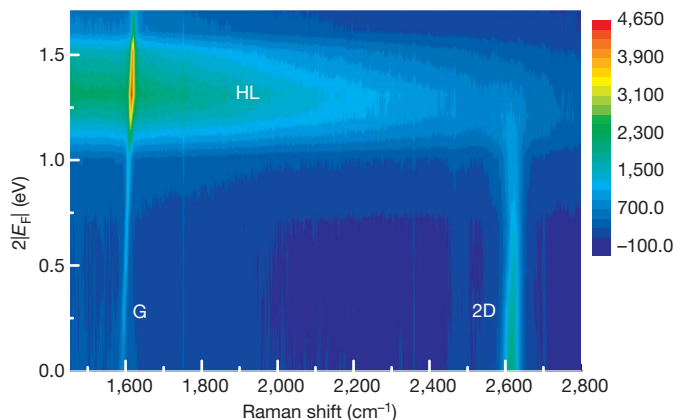


Figure 2 | Controlling inelastic light scattering with electrostatic doping. Graphene inelastic light scattering intensity (colour scale, arbitrary units) is plotted as a function of Stokes Raman redshift and $2|E_F|$ for 1.58-eV laser excitation energy (E_{ex}). G-mode Raman intensity (G; $\sim 1,600 \text{ cm}^{-1}$) is strongly enhanced as $2|E_F|$ approaches E_{ex} , thus blocking some of the Raman quantum pathways. In contrast, the 2D-Raman peak (2D; $\sim 2,600 \text{ cm}^{-1}$) is suppressed and eventually disappears with increasing $2|E_F|$. Broadband hot luminescence (HL) emerges when $2|E_F|$ is lower than but close to E_{ex} .

with reduced optical conductivity below the threshold energy, $2|E_F|$. The transition width of the threshold is due largely to the lifetime broadening from the excited state, and the transition centre yields the $2|E_F|$ value with an uncertainty less than 100 meV. From these spectra, we can also determine the carrier concentration $n = (E_F/\hbar v_F)^2/\pi$ (refs 10, 11), where v_F is the Fermi velocity. With ion-gel gating, vertical electronic transitions with excitation energies as high as 1.8 eV can be blocked. This corresponds to an induced carrier concentration of $6 \times 10^{13} \text{ cm}^{-2}$.

This large gate-induced shift in Fermi energy with ion-gel gating not only allows for controlled optical absorption, but also enables control over inelastic light scattering by varying the allowed intermediate excitations. Using a 785-nm excitation laser (that is, $E_{\text{ex}} = 1.58 \text{ eV}$),

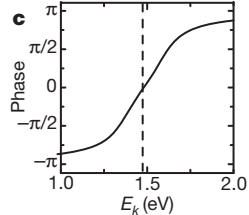
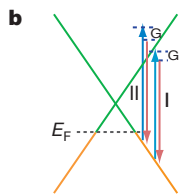
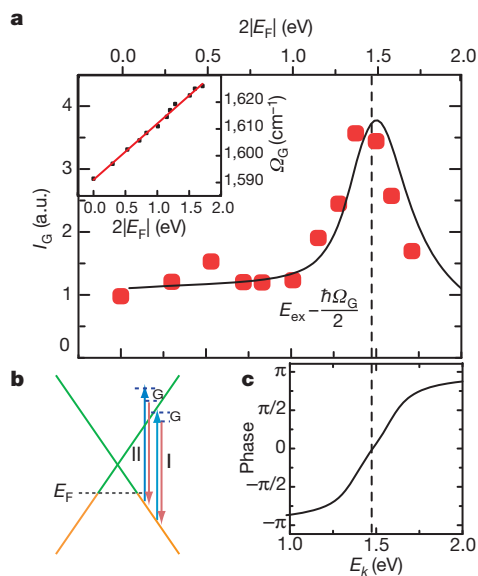
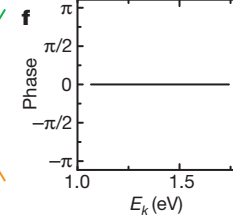
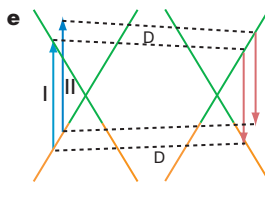
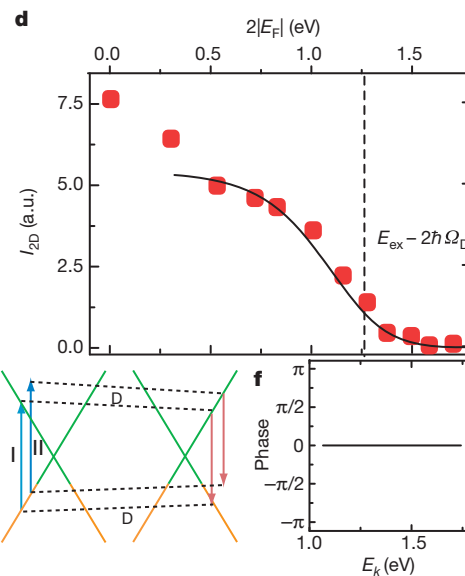


Figure 3 | Quantum interference between graphene Raman pathways. **a**, G-mode Raman intensity as a function of $2|E_F|$ (symbols). A large enhancement is observed when some of the Raman pathways are blocked, which is a hallmark of quantum interference. Black line, theoretical prediction. Inset, doping dependence of G-mode frequency. **b**, Illustration of two representative G-mode Raman pathways (I and II) through two different intermediate excited states. **c**, Quantum phase of Raman pathway amplitude illustrated as a function of intermediate excitation energy. Raman pathways with intermediate excitation energy above and below $E_{\text{ex}} - \hbar\Omega_G/2$ (dashed

we examine the evolution of inelastic light scattering spectra as $2|E_F|$ is varied from 0 to 1.8 eV. Figure 2 shows the inelastic emission intensity as a function of the Raman shift and of $2|E_F|$. At low doping, the spectra are characterized by prominent G-mode ($\sim 1,600 \text{ cm}^{-1}$) and 2D-mode ($\sim 2,600 \text{ cm}^{-1}$) phonon Raman peaks, as observed in previous studies¹². But two new features emerge when $2|E_F|$ becomes large enough to significantly alter intermediate-state pathways. (1) With larger $2|E_F|$ blocking part of the Raman quantum pathways, the G-mode Raman intensity does not decrease as one might expect, but increases and reaches a much higher peak value. This behaviour is in contrast to 2D-mode Raman intensity, the intensity of which does decrease monotonically as the resonant Raman pathways are blocked. (2) Photoluminescence over a wide spectral range emerges when $2|E_F|$ is slightly below the laser excitation energy. This hot-electron luminescence from continuous-wave laser excitation is quite unexpected, and it is distinctly different from recently observed nonlinear photoluminescence using femtosecond laser excitations^{23,24}.

We start our investigation by examining the unusual behaviour of G-mode phonon Raman intensity. The measured Raman peak position Ω_G blue-shifts linearly with Fermi energy as $\Delta\Omega_G = |E_F| \times 42 \text{ cm}^{-1} \text{ eV}^{-1}$ (Fig. 3a inset). This blue shift is due to phonon renormalization from non-adiabatic electron-phonon coupling, and has been well-studied previously^{16,17}. The behaviour of the integrated G-mode Raman intensity with $2|E_F|$ (Fig. 3a), however, has not been observed before. Increased event probability with reduced pathways, as observed in the G-mode Raman intensity here, is a canonical signature of destructive quantum interference. This result highlights the quantum nature of Raman scattering, and provides a valuable probe for microscopic Raman processes in graphene.

G-phonon Raman scattering has been extensively studied for characterizing graphitic materials^{12,13,25}, and is widely assumed to be a resonant Raman scattering process dominated by pathway I depicted in Fig. 3b (ref. 13). This picture, although appealing, cannot account for our observation. The possibility of enhancement in the Raman G signal on doping because of quantum interference between different Raman pathways was discussed in a recent theoretical study⁸. The



line) have an average phase difference of π and interfere destructively. Blocking pathways below $E_{\text{ex}} - \hbar\Omega_G/2$ therefore leads to large enhancement of G-mode Raman scattering. **d**, Doping dependence of 2D-mode Raman intensity (symbols). It drops quickly to zero when $2|E_F|$ approaches $E_{\text{ex}} - 2\hbar\Omega_D$ (dashed line), and agrees well with two-phonon Raman scattering theory (black line). **e**, Two representative pathways for the 2D Raman mode. **f**, Two-phonon scattering pathways illustrated in **e** have the same phase and interfere constructively. Blocking these pathways diminishes the intensity of 2D Raman scattering.

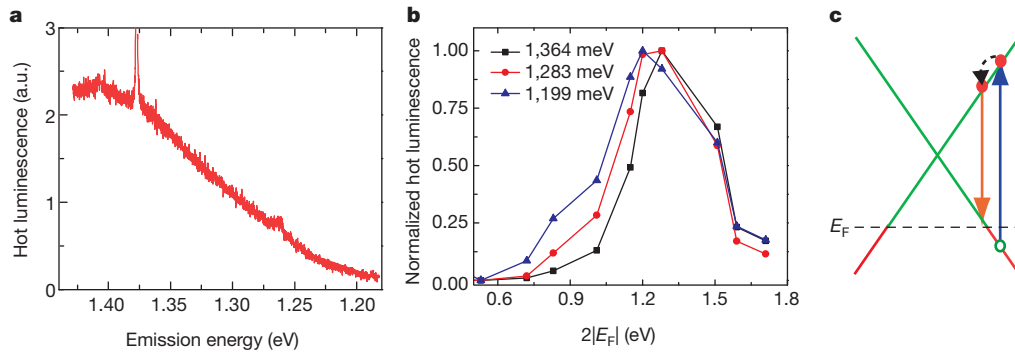


Figure 4 | Hot luminescence in graphene. **a**, Hot luminescence at $2|E_F| = 1.4$ eV shows a broad emission spectrum. It has an integrated intensity (in the range 1.2–1.4 eV) more than 100 times stronger than phonon Raman scattering. **b**, Normalized luminescence intensity as a function of $2|E_F|$ for three different excitation energies. Higher energy photon emissions have higher onset values for $2|E_F|$. **c**, Illustration of the hot luminescence generation mechanism. Red and green

Raman pathways in graphene can be on resonance (I in Fig. 3b), close to resonance (II in Fig. 3b), or off resonance. For a given final state of excitation (that is, emission of an optical phonon with defined momentum and polarization), pathways involving different intermediate states are indistinguishable. Therefore, all the pathways having different quantum mechanical amplitudes (that is, phase and magnitude) will interfere with each other. As such, blocking one quantum pathway can not only decrease but also increase the Raman intensity, depending on its phase relative to other pathways.

G-mode Raman intensity for any specific final one-phonon and one-photon state can be described by⁶:

$$I = \left| \sum_k C_k R_k \right|^2$$

and

$$R_k = \frac{1}{(E_{\text{ex}} - E_k - i\gamma)(E_{\text{ex}} - \hbar\Omega_G - E_k - i\gamma)}. \quad (1)$$

Here C_k and R_k are the matrix element and resonance factor, respectively, for a Raman pathway through a vertical electronic transition at wavevector k . E_k is the transition energy, E_{ex} the excitation energy, Ω_G the G-phonon frequency, and γ the energy broadening of the excited state. Phase differences between different pathways come mainly from the resonance factor R_k . For example, R_k for pathways resonant with incident light ($E_k = E_{\text{ex}}$) and scattered light ($E_k = E_{\text{ex}} - \hbar\Omega_G$) have opposite sign for small γ , and their Raman amplitudes largely cancel one other. In Fig. 3c we plot the phase of quantum pathways, $\Phi = \arg(R_k)$, through different intermediate electronic transitions for $E_{\text{ex}} = 1.58$ eV. We have set $\gamma = 0.2$ eV, which is estimated from the energy broadening in optical absorption spectra. The phase varies rapidly close to the resonance condition $E_k \approx E_{\text{ex}}$. In particular, pathways with transition energy E_k above and below energy $E_{\text{ex}} - \hbar\Omega_G/2$ have an average phase difference of π . They interfere destructively if all quantum pathways are allowed, and this leads to a weak overall Raman signal. When $2|E_F|$ is increased to block pathways with transition energies below $E_{\text{ex}} - \hbar\Omega_G/2$, the destructive interference is eliminated, and this leads to the observed large enhancement of Raman intensity. When $2|E_F|$ is increased further, more in-phase pathways are blocked and the Raman intensity starts to decrease. A quantitative evaluation of the matrix elements and a summation of all quantum pathways with $\gamma = 0.2$ eV was performed following ref. 8, and the result (black line in Fig. 3a) fits our experimental data (red squares) well. Because the cancellation from destructive interference from different Raman pathways is not perfect with energy dependent transition matrix elements and density of states, the G-mode Raman signal is still observable in undoped graphene. Nevertheless, the Raman intensity with all resonant pathways blocked ($2|E_F| = 1.8$ eV) is approximately the same

as when every pathway is allowed ($2|E_F| = 0$). It shows that the simple picture of G-mode graphene Raman scattering, in which only the on-resonance pathways are considered, is invalid.

The integrated intensity of the 2D-mode Raman peak, which results from simultaneous excitation of two phonons with wavevector close to the K-point in the graphene Brillouin zone, exhibits a doping dependence completely different from that of the integrated intensity of the G-mode Raman peak. Figure 3d shows the 2D-mode Raman intensity as a function of $2|E_F|$. The 2D intensity decreases slowly with increased doping when $2|E_F|$ is relatively small, which was observed previously¹⁴ and can be explained theoretically through the excited state broadening^{14,26}. Here we focus on the behaviour when $2|E_F|$ approaches E_{ex} . Instead of an enhanced scattering intensity as in G-mode Raman, the 2D-mode intensity drops quickly at the threshold energy $2|E_F| = E_{\text{ex}} - 2\hbar\Omega_D$, where Ω_D is the zone-boundary phonon frequency.

The intensity of 2D-mode Raman scattering is also determined by quantum interference between different pathways. Figure 3e illustrates two representative Raman pathways, including one satisfying the double resonance condition (pathway I)¹². Such two-phonon Raman pathways have one more virtual transition than one-phonon processes, and it has a profound effect on the Raman interference behaviour. Quantitatively, quantum interference in 2D-mode Raman for a specific final two-phonon and one-photon state is described by⁶:

$$I = \left| \sum_k D_k R_k \right|^2, \quad (2)$$

$$R_k = \frac{1}{(E_{\text{ex}} - E_k - i\gamma)(E_{\text{ex}} - \hbar\Omega_D - E_k - i\gamma)(E_{\text{ex}} - 2\hbar\Omega_D - E_{k'} - i\gamma)}$$

Here D_k is the matrix element of 2D-mode Raman scattering for a pathway starting at transition k . E_k , $E_{k'}$ and $E_{k''}$ are the energies of the three intermediate states. The resonance factor R_k determines the relative phase between different pathways. Figure 3f plots the phase of quantum pathways, involving the emission of two phonons with a specific wavevector, as a function of first-intermediate-state energy E_k for processes similar to those depicted in Fig. 3e (see also Supplementary Information). All the different pathways have the same phase and they interfere constructively. As a result, any reduction in quantum pathways leads to decreased Raman intensity, and when all resonant channels are blocked, the 2D-mode Raman signal becomes negligible. In Fig. 3d we show a more quantitative comparison between this theoretical picture (black line) and experimental results (symbols).

Last, we examine the graphene hot luminescence. Figure 4a shows the hot luminescence spectrum at $2|E_F| = 1.4$ eV, which is characterized by a broadband emission. To investigate the effect of carrier doping, we plot in Fig. 4b the normalized luminescence intensity as a function of

$2|E_F|$ for several emission energies (symbols). The luminescence is distinctly different from electronic Raman scattering²⁷ because zero-momentum electron-hole pair excitations required for the electronic Raman scattering do not exist in the strongly doped graphene. The peak observed in the hot luminescence when $2|E_F|$ approaches E_{ex} also has a different physical origin compared with that in G-mode phonon Raman. Instead of being an interference phenomenon from blocked Raman quantum pathways, the hot luminescence maximum is a consequence of newly opened dynamic pathways for photo-excited hot carriers.

We depict this dynamic process in Fig. 4c. An incident photon generates an electron and a hole, which then relax to the Fermi level through interactions with low-energy electrons and phonons. During the relaxation process, a hot electron has a finite probability of emitting a photon, but requires the final valence band state to be empty because of Pauli blocking (Fig. 4c). This pathway is opened up only when $2|E_F|$ exceeds the light emission energy, and defines a threshold doping level for hot luminescence generation. The qualitative features observed in Fig. 4b can be understood using this picture: hot luminescence emerges when the energy $2|E_F|$ is higher than an onset value that increases with the emission photon energy. When $2|E_F|$ is further increased to a value greater than the laser energy, hot luminescence at all energies suddenly disappears as the initial photoexcitations are blocked. Because the hot luminescence arises from radiative decay of excited electrons right after photoexcitation, it could become a valuable probe for ultrafast electron dynamics in graphene⁹.

Our study shows that inelastic light scattering phenomena can be explored by controlling the intermediate excited states. In graphene, it enables us to demonstrate the critical role of quantum interference in Raman scattering, and to reveal a new broadband hot luminescence. Such inelastic light scattering control could also be applied to general nanoscale material research, because electronic transitions in many nanostructures can be modified in a controlled fashion (for example, electrostatic gating of carbon nanotubes²⁸ and electrochemically tuned semiconductor quantum dots²⁹). This control will make inelastic light scattering a more powerful tool for probing novel nanoscale physics. It could also lead to optimized inelastic light scattering in nanomaterials for biological sensors and optoelectronic applications.

METHODS SUMMARY

In this study we use large area graphene grown by chemical vapour deposition³⁰. Graphene is grown on copper films using CH_4 as the feed gas, which was then transferred with poly(methyl methacrylate) (PMMA) support to a fused silica substrate after wet-etching to remove the copper film by $FeCl_3$. The PMMA support is dissolved in acetone solution. Subsequently, Ti (10 nm) and Au (40 nm) were deposited in vacuum through stencil masks onto the graphene sample for the source, drain and gate electrodes.

The ionic liquid, 1-ethyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide ([EMIM][TFSI]), was purchased from EMD Chemicals. The ionic liquid was dried at $T = 100^\circ C$ under vacuum (~ 200 mtorr) for 3 days and then transferred to an inert-atmosphere glove box. Polystyrene-*b*-poly(ethylene oxide)-*b*-polystyrene (PS-PEO-PS) triblock copolymer was purchased from Polymer Source and used as received. The quoted molecular weights of the block copolymer moieties were 10–44–10 kg mol⁻¹ for the PS-PEO-PS blocks, respectively (PEO volume fraction = 0.67). The physically crosslinked, ionic liquid/triblock copolymer gel (ion gel) was produced in the following manner. In an inert atmosphere glove box, 0.55 g of [EMIM][TFSI] were dissolved with 21 mg of PS-PEO-PS in 2 ml of dry dichloromethane. The solution was stirred overnight at room temperature. The ion gel was removed from the glove box and spin-coated on the graphene sample at a rate of 4,000 r.p.m.

Infrared transmission measurements were performed with a Fourier transform infrared spectrometer at the Advanced Light Source, Lawrence Berkeley National Laboratory. Inelastic light scattering was measured with a micro-Raman set-up with a 785-nm excitation laser. The Raman set-up has a spectral resolution of 1 cm⁻¹. All measurements were performed in air and at 300 K.

Received 1 September 2010; accepted 18 January 2011.

Published online 16 March 2011.

1. Raman, C. V. A change of wave-length in light scattering. *Nature* **121**, 619 (1928).

2. Landsberg, G. & Mandelstam, L. Eine neue Erscheinung bei der Lichtzerstreuung in Kristallen. *Naturwissenschaften* **16**, 557 (1928).
3. Rao, A. M. *et al.* Diameter-selective Raman scattering from vibrational modes in carbon nanotubes. *Science* **275**, 187–191 (1997).
4. Devereaux, T. P. & Hackl, R. Inelastic light scattering from correlated electrons. *Rev. Mod. Phys.* **79**, 175–233 (2007).
5. Gofri, A. R. *et al.* One-dimensional plasmon dispersion and dispersionless intersubband excitations in GaAs quantum wires. *Phys. Rev. Lett.* **67**, 3298–3301 (1991).
6. Cardona, M. *Light Scattering in Solids I* 2nd edn (Springer, 1982).
7. Ralston, J. M., Wadsack, R. L. & Chang, R. K. Resonant cancellation of Raman scattering from CdS and Si. *Phys. Rev. Lett.* **25**, 814–818 (1970).
8. Basko, D. M. Calculation of the Raman G peak intensity in monolayer graphene: role of Ward identities. *N. J. Phys.* **11**, 095011 (2009).
9. Elsaesser, T., Shah, J., Rota, L. & Lugli, P. Initial thermalization of photoexcited carriers in GaAs studied by femtosecond luminescence spectroscopy. *Phys. Rev. Lett.* **66**, 1757–1760 (1991).
10. Novoselov, K. S. *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197–200 (2005).
11. Zhang, Y. B., Tan, Y. W., Stormer, H. L. & Kim, P. Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature* **438**, 201–204 (2005).
12. Ferrari, A. C. *et al.* Raman spectrum of graphene and graphene layers. *Phys. Rev. Lett.* **97**, 187401 (2006).
13. Pimenta, M. A. *et al.* Studying disorder in graphite-based systems by Raman spectroscopy. *Phys. Chem. Chem. Phys.* **9**, 1276–1291 (2007).
14. Das, A. *et al.* Monitoring dopants by Raman scattering in an electrochemically top-gated graphene transistor. *Nature Nanotechnol.* **3**, 210–215 (2008).
15. Dresselhaus, M. S., Jorio, A., Hofmann, M., Dresselhaus, G. & Saito, R. Perspectives on carbon nanotubes and graphene Raman spectroscopy. *Nano Lett.* **10**, 751–758 (2010).
16. Pisana, S. *et al.* Breakdown of the adiabatic Born-Oppenheimer approximation in graphene. *Nature Mater.* **6**, 198–201 (2007).
17. Yan, J., Zhang, Y. B., Kim, P. & Pinczuk, A. Electric field effect tuning of electron-phonon coupling in graphene. *Phys. Rev. Lett.* **98**, 166802 (2007).
18. Li, Z. Q. *et al.* Dirac charge dynamics in graphene by infrared spectroscopy. *Nature Phys.* **4**, 532–535 (2008).
19. Wang, F. *et al.* Gate-variable optical transitions in graphene. *Science* **320**, 206–209 (2008).
20. Zhang, Y. *et al.* Direct observation of a widely tunable bandgap in bilayer graphene. *Nature* **459**, 820–823 (2009).
21. Cho, J. H. *et al.* Printable ion-gel gate dielectrics for low-voltage polymer thin-film transistors on plastic. *Nature Mater.* **7**, 900–906 (2008).
22. Kim, B. J. *et al.* High-performance flexible graphene field effect transistors with ion gel gate dielectrics. *Nano Lett.* **10**, 3464–3466 (2010).
23. Lui, C. H., Mak, K. F., Shan, J. & Heinz, T. F. Ultrafast photoluminescence from graphene. Preprint at (<http://arXiv.org/abs/1006.5769>) (2010).
24. Stoehr, R. J., Kolesov, R., Pflaum, J. & Wrachtrup, J. Fluorescence of laser created electron-hole plasma in graphene. Preprint at (<http://arXiv.org/abs/1006.5434>) (2010).
25. Dresselhaus, M. S., Dresselhaus, G., Saito, R. & Jorio, A. Raman spectroscopy of carbon nanotubes. *Phys. Rep.* **409**, 47–99 (2005).
26. Basko, D. M., Pisanec, S. & Ferrari, A. C. Electron-electron interactions and doping dependence of the two-phonon Raman intensity in graphene. *Phys. Rev. B* **80**, 165413 (2009).
27. Kashuba, O. & Fal'ko, V. I. Signature of electronic excitations in the Raman spectrum of graphene. *Phys. Rev. B* **80**, 241404(R) (2009).
28. Ilani, S., Donev, L. A. K., Kindermann, M. & McEuen, P. L. Measurement of the quantum capacitance of interacting electrons in carbon nanotubes. *Nature Phys.* **2**, 687–691 (2006).
29. Wang, C. J., Shim, M. & Guyot-Sionnest, P. Electrochromic nanocrystal quantum dots. *Science* **291**, 2390–2392 (2001).
30. Li, X. S. *et al.* Large-area synthesis of high-quality and uniform graphene films on copper foils. *Science* **324**, 1312–1314 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the US Department of Energy, Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory under contract no. DE-AC02-05CH11231 (C.-F.C. and F.W.), by the Office of Basic Energy Sciences under contract nos DE-AC02-05CH11231 (B.W.B. and R.A.S.), DE-AC03-76SF0098 (Materials Science Division) (C.G., A.Z.) and DE-AC02-05CH11231 (Advanced Light Source), and by ONR MURI award N00014-09-1-1066 (J.H., C.-H.P., S.G.L., M.F.C.). C.-F.C. also acknowledges fellowship support from the National Science Council and National Tsing Hua University, Taiwan, under awards NSC98-2811-M-007-008 and NSC98-2120-M-007-004.

Author Contributions F.W. designed the experiment; C.-F.C. and J.H. carried out optical measurements; B.G., C.G. and B.W.B. contributed to sample growth and fabrication; and C.-H.P., S.G.L. and F.W. performed theoretical analysis. All authors discussed the results and wrote the paper together.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to F.W. (fengwang76@berkeley.edu).

Structure of mammalian AMPK and its regulation by ADP

Bing Xiao^{1*}, Matthew J. Sanders^{2†*}, Elizabeth Underwood^{1*}, Richard Heath^{1*}, Faith V. Mayer^{2*}, David Carmena², Chun Jing¹, Philip A. Walker¹, John F. Eccleston^{1‡}, Lesley F. Haire¹, Peter Saiu^{1†}, Steven A. Howell¹, Rein Aasland^{1†}, Stephen R. Martin¹, David Carling² & Steven J. Gambin¹

The heterotrimeric AMP-activated protein kinase (AMPK) has a key role in regulating cellular energy metabolism; in response to a fall in intracellular ATP levels it activates energy-producing pathways and inhibits energy-consuming processes¹. AMPK has been implicated in a number of diseases related to energy metabolism including type 2 diabetes, obesity and, most recently, cancer^{2–6}. AMPK is converted from an inactive form to a catalytically competent form by phosphorylation of the activation loop within the kinase domain⁷: AMP binding to the γ -regulatory domain promotes phosphorylation by the upstream kinase⁸, protects the enzyme against dephosphorylation, as well as causing allosteric activation⁹. Here we show that ADP binding to just one of the two exchangeable AXP (AMP/ADP/ATP) binding sites on the regulatory domain protects the enzyme from dephosphorylation, although it does not lead to allosteric activation. Our studies show that active mammalian AMPK displays significantly tighter binding to ADP than to Mg-ATP, explaining how the enzyme is regulated under physiological conditions where the concentration of Mg-ATP is higher than that of ADP and much higher than that of AMP. We have determined the crystal structure of an active AMPK complex. The structure shows how the activation loop of the kinase domain is stabilized by the regulatory domain and how the kinase linker region interacts with the regulatory nucleotide-binding site that mediates protection against dephosphorylation. From our biochemical and structural data we develop a model for how the energy status of a cell regulates AMPK activity.

At the whole-body level, AMPK is regulated by a diverse range of hormones—for example, leptin¹⁰, adiponectin¹¹, ciliary neurotrophic factor¹² and ghrelin¹³—and it has a role in appetite^{13,14}, glucose, lipid and protein metabolism^{1,3,15}, cell growth, and cell polarity^{2,4}. AMPK is a heterotrimeric complex comprising an α -catalytic subunit and two regulatory subunits (β and γ)¹. Activation of AMPK requires phosphorylation of Thr 172, which lies in the activation segment of the amino-terminal kinase domain of the α -subunit⁷. Phosphorylation of Thr 172 leads to a several-hundred-fold increase in activity^{9,16}. In mammals, calcium/calmodulin-dependent protein kinase kinase- β (CaMKK β) and LKB1 are the predominant kinases upstream of AMPK, although there is some evidence implicating other upstream kinases^{17,18}. Previous studies have shown that AMP protects against the dephosphorylation of Thr 172 (refs 16, 19) and we recently provided evidence that protection against dephosphorylation is the major physiological mechanism for activation of AMPK⁹. In addition to activation by phosphorylation, AMP causes a 2–5-fold allosteric activation of AMPK depending on the nature of the isoforms present in the AMPK complex²⁰. To address this issue we previously investigated the nucleotide-binding properties of the γ 1 subunit of AMPK and determined the structure of the regulatory core of mammalian

AMPK (C-terminal domain of α -subunit, C-terminal domain of β -subunit, full-length γ -subunit; hereafter referred to as the regulatory fragment) in complex with nucleotides²¹. Notably, our studies revealed that three of the four potential nucleotide-binding sites are occupied²¹. One of these sites contains a permanently bound AMP molecule (site 4, following nomenclature proposed by ref. 22), whereas AMP and Mg-ATP compete for binding at the other two sites²¹ (site 1 and site 3).

Unlike AMP, ADP has no significant allosteric effect on AMPK isolated from rat liver²³. Consistent with this, we also found that ADP does not activate recombinant AMPK under conditions where AMP produces a twofold activation (Fig. 1a). However, our studies show that ADP provides protection of AMPK from dephosphorylation across a similar range of concentrations as AMP (Fig. 1b). We have also shown the same effect using AMPK purified from rat liver (Supplementary Fig. 2a). Although Mg-ATP does not protect against dephosphorylation (Fig. 1c), it does compete with the protective effect of both AMP and ADP on dephosphorylation (Fig. 1d). We have also shown that the protective effect of ADP is lost in a Wolff–Parkinson–White syndrome mutation (Supplementary Fig. 2b). We propose that AMP or ADP (AMP/ADP) binding shifts the equilibrium between dephosphorylation-sensitive and -insensitive states, and thus slows, but does not abolish, dephosphorylation of the enzyme by phosphatases.

Extending our earlier work looking at nucleotide binding to the regulatory fragment, we characterized binding of nucleotides to active full-length AMPK. For these studies we used CaMKK β to stoichiometrically phosphorylate Thr 172 on the activation loop of recombinant full-length AMPK, we used the coumarin adducts of ATP and ADP as fluorescent reporters of nucleotide binding, and derived the binding parameters for the unlabelled nucleotides by competition experiments (Fig. 2a). We verified that the two species bind at the same sites by determining the crystal structures of the regulatory fragment in complex with coumarin-ADP and with ADP (Supplementary Fig. 3). The results show (Table 1) that the two exchangeable sites have markedly different affinities for nucleotides. Binding at the tighter of the two sites is at least 30-fold stronger than at the weaker site. Given that under physiological conditions most of the ATP is coordinated to Mg²⁺, and that the majority of AMP and ADP is not, we also measured nucleotide binding in the presence of this cation. The data show that Mg-ATP binds up to tenfold weaker than ATP. Thus, active AMPK binds AMP/ADP significantly more strongly than it does Mg-ATP at both exchangeable sites. There are two lines of evidence that lead us to conclude that it is AMP/ADP binding at the weaker of the two exchangeable sites that accounts for the protection of the enzyme against dephosphorylation. The first is that the dose-response curve for AMP/ADP-mediated protection against phosphorylation correlates with the binding curves for these nucleotides at the weaker, rather than the stronger, of the two binding sites (Fig. 1b). The second comes from our discovery that NADH binds to AMPK.

¹MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK. ²MRC Clinical Sciences Centre, Hammersmith Hospital Campus, Imperial College, DuCane Road, London W12 0NN, UK. [†]Present addresses: MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK (M.J.S.); Department of Molecular Biology, University of Bergen, Thormøhlensgt. 55, N5020 Bergen, Norway (R.A.); Cancer Research UK, 44 Lincoln's Inn Fields, PO Box 123, London WC2A 3PX, UK (P.S.).

*These authors contributed equally to this work.

‡Deceased.

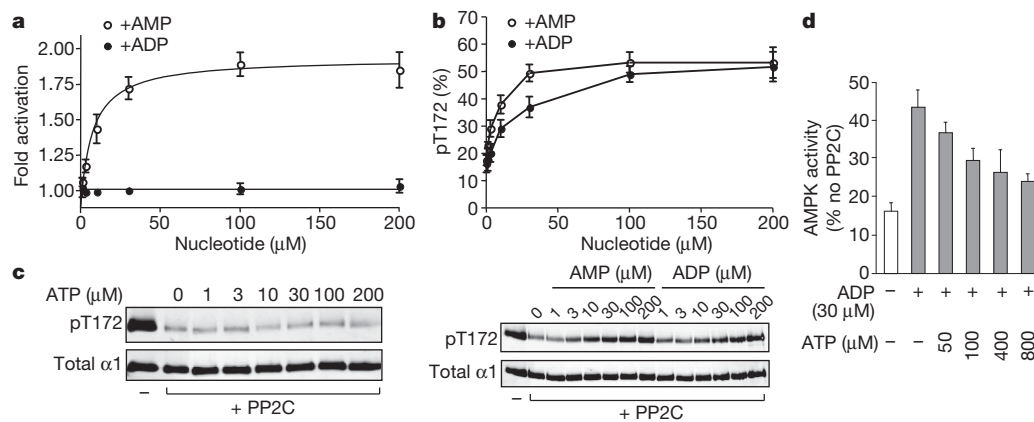


Figure 1 | Role of ADP in regulation of AMPK activity. **a**, AMP, but not ADP, allosterically activates AMPK. **b**, AMP and ADP protect AMPK from dephosphorylation. **c**, ATP does not protect against dephosphorylation. **d**, Mg-ATP competes with the protective effect of ADP on dephosphorylation. All results are displayed as the mean \pm s.e.m. determined from at least three independent experiments. Where appropriate a representative blot ($n = 3$) showing Thr 172 phosphorylation and total α -subunit levels is shown.

NADH undergoes a significant change in fluorescence upon binding to AMPK. We used this property to establish that the cofactor binds to a single site on the enzyme, with a dissociation constant (K_d) of about 50 μM (Fig. 2b, inset). NADH binding is competed by AXPs binding to the stronger, but not the weaker, of the two exchangeable sites (Fig. 2b and Table 1). When we repeated the ADP protection against dephosphorylation experiments using a range of NADH concentrations, we found no evidence for NADH competing with the protective effect of ADP on dephosphorylation, whereas NADH and ADP both compete with AMP for allosteric activation of the enzyme (Supplementary Fig. 4). This observation indicates that it is AMP/ADP binding at the weaker of the two exchangeable sites, the one that does not bind NADH, that is responsible for protection against dephosphorylation. We also carried out co-crystallization of the regulatory fragment with one molar equivalent of ADP (Supplementary Fig. 5). The resulting electron density map showed full occupancy of ADP at site 1 and no detectable density at site 3, identifying site 3 as the weaker binding site. We can therefore assign the allosteric effect to AMP binding at the tighter site 1, and protection against dephosphorylation is mediated by AMP/ADP binding at the weaker site 3.

Previous studies on the regulation of AMPK have focused on the role of AMP because it allosterically activates the enzyme¹⁵ whereas ADP does not. However, phosphorylation remains central to AMPK regulation as the enzyme is inactive in the absence of Thr 172 phosphorylation^{7,16}. Under optimal conditions, mammalian cells maintain ATP at a high level relative to ADP and AMP. Typical concentrations

of free adenine nucleotides in mammalian cells lie in the range of 3,000–8,000 μM for ATP, 50–200 μM for ADP and 0.5–5 μM for AMP^{24–26}. Because the free concentration of ADP is between 10- to 400-fold higher than AMP, and their binding constants are similar, ADP will be more successful at competing with Mg-ATP than AMP. Therefore, the fact that ADP protects AMPK from dephosphorylation is likely to represent an important physiological mechanism for regulating the activity of the enzyme.

We have also determined the crystal structure of an active form of the enzyme that encompasses the whole of the catalytic α -subunit. The construct used is shown in Fig. 3a (details of its design are given in Supplementary Fig. 6). From the best samples we collected a data set to 3.3 \AA Bragg spacing, after screening about 100 crystals, and solved the structure by molecular replacement using independent models for the regulatory fragment (Protein Data Bank 2V8Q)²¹ and the kinase domain (Protein Data Bank 2H6D)²⁷. Although the data set is at medium resolution, the molecular replacement solution was robust and yielded initial electron density that revealed the location of many components that were not present in the original model. As might be expected for a structure of this complexity and resolution, some parts of the molecule are better defined than others. For example, the activation loop of the kinase domain, which is packed against the regulatory fragment, has better defined electron density than loops on the surface of the complex, which often show continuous main-chain density but lack side-chain features. Overall, the most important features of the current structure concern the architecture of the complex, particularly the relationships between the α -kinase domain and the α -linker with the regulatory fragment (omit maps for these regions are presented in Supplementary Fig. 7).

The structure is shown in Ribbons representation (Fig. 3b, c and Supplementary Fig. 8) and in space-filling representation before and after the kinase domain and the linker region have been rotated away from the complex to display their contact regions (Fig. 3d). The first of these interfaces involves the kinase domain and has a surface area of about 1,100 \AA^2 . This relatively modest contact area is consistent with the observation that specific protease cleavage of the linker between the kinase domain and the C-terminal domain of the α -subunit leads to material that separates into two components (kinase domain plus regulatory fragment) after gel filtration. A significant part of the

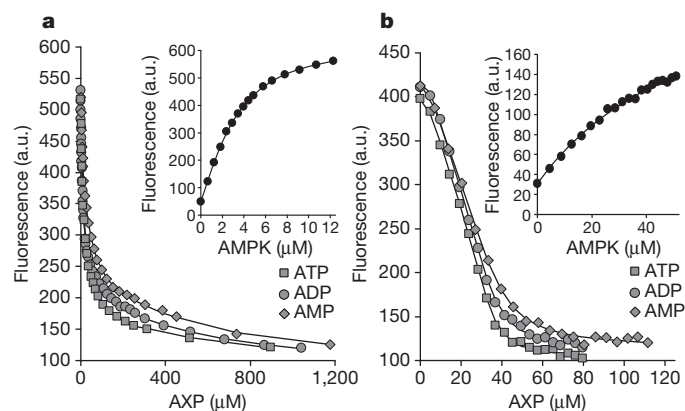


Figure 2 | Measurement of equilibrium dissociation constants for the binding of AXPs to phosphorylated AMPK. **a**, Displacement of coumarin-ATP (C-ATP) from the AMPK-(coumarin-ATP)₂ complex by AXPs monitored using fluorescence at 470 nm. Solid lines are the computed best fits with $K_{d,I}$ and $K_{d,II}$ for C-ATP binding to AMPK fixed at 1.1 and 4.2 μM . Inset: titration of coumarin-ATP with AMPK. **b**, Displacement of NADH from the AMPK-NADH complex by AXPs monitored using fluorescence at 435 nm. The solid line is the computed best fit with the K_d for NADH fixed at 65 μM . Inset: fluorescence titration of NADH with AMPK.

Table 1 | Equilibrium K_d values for the binding of AXPs to phosphorylated AMPK

Ligand	K_d (μM) versus NADH	$K_{d,I}$ (μM)	$K_{d,II}$ (μM)
		versus C-AXPs	
AMP	1.6 (0.5)	2.5 (0.6)	80 (25)
ADP	1.3 (0.5)	1.5 (0.4)	50 (15)
ATP	0.9 (0.3)	1.7 (0.5)	65 (15)
Mg-ATP	32 (12)	18 (7.5)	230 (80)

Dissociation constants (K_d) were determined at 20 $^{\circ}\text{C}$ by competition against NADH or coumarin-ATP (C-AXPs) in 25 mM Tris, 1 mM TCEP, 100 mM NaCl (pH 8) with and without 5 mM MgCl_2 . The K_d values are reported as the mean (\pm s.d.) determined from at least three independent measurements.

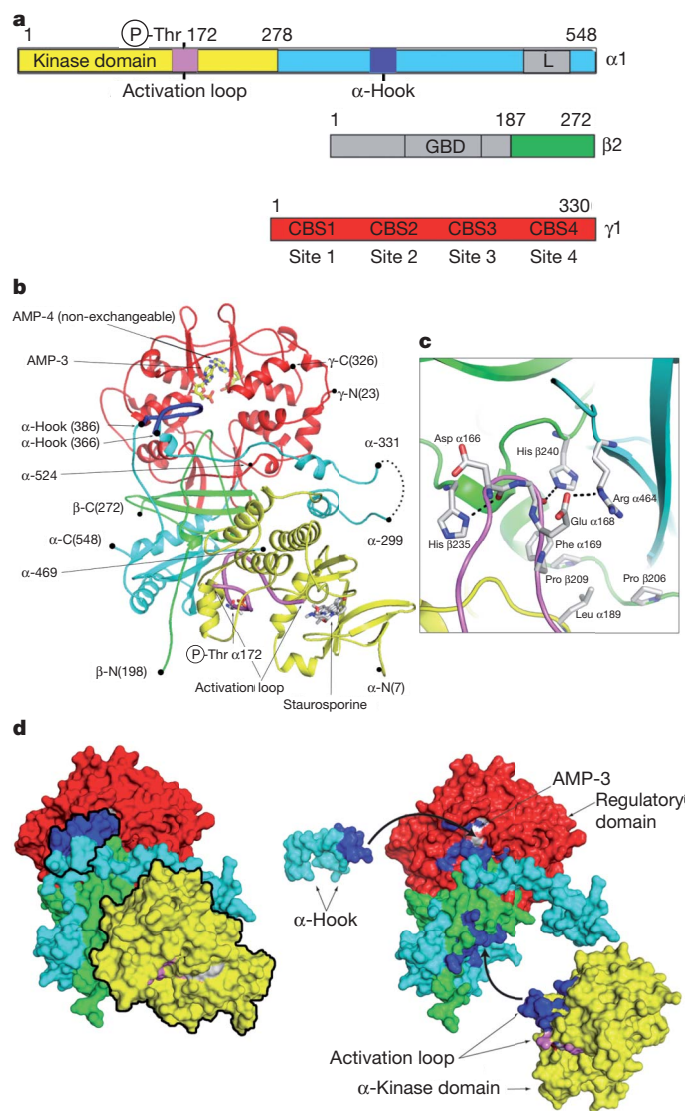


Figure 3 | Crystal structure of active mammalian AMPK. **a**, Schematic representation of the components of the heterotrimer; the parts of the complex missing from the crystallized protein are shown in grey. The domains, including the activation loop (pink) and α -hook (dark blue), are coloured the same in all panels. GBD, glycogen-binding domain. **b**, Ribbon representation of the crystallized complex with two bound AMPs, staurosporine and phosphorylated Thr 172 shown in stick representation. The α -hook and activation loop of the kinase domain are shown in heavier lines and coloured dark blue and pink, respectively. **c**, The interface between the activation loop and the regulatory fragment is shown in more detail in a similar orientation as **b**; potential electrostatic interactions are indicated by dashed lines. **d**, The complex is shown in two space-filling representations. The left panel represents the same view as **b** with the α -hook and kinase domain outlined in black. In the right-hand panel these two components have been rotated away from the regulatory domain to show the interaction surfaces in an 'open-book' representation, where the contacting residues have been coloured in dark blue. With the α -hook removed, AMP-3 becomes visible. The black arrows indicate the rotations that would reassemble the complex.

contact area involves the activation loop of the kinase interacting predominantly with the C terminus of the β -domain (Fig. 3c). Unlike previous structures of the isolated and non-phosphorylated kinase domain, the phosphorylated activation loop in our structure is well ordered, as evidenced by electron density maps (Supplementary Fig. 7). The small lobe of the kinase is in a more closed conformation relative to the unphosphorylated (and thus inactive) isolated kinase domain structure²⁷ (Supplementary Fig. 9). The fact that the activation loop mediates the interaction of the kinase domain with the regulatory

fragment means that, in this conformation, Thr 172 is protected from access by phosphatases. This idea is strongly supported by site-directed mutagenesis experiments: mutation of β 1 His 233 (corresponding to His-235 in β 2) at the interface with the kinase domain (Fig. 3c) results in an enzyme that is activatable by phosphorylation but that has a significantly increased rate of dephosphorylation in phosphatase assays (Fig. 4a, b).

Another component of the α -subunit-regulatory fragment interaction is provided by a part of the segment of the α -chain that links the N-terminal kinase domain to the C-terminal regulatory fragment, involving residues between α 373 and α 382 that are largely conserved between α 1 and α 2 in vertebrates (Supplementary Fig. 10) and which we have called the α -hook structure (Fig. 3b, d). The α -hook interacts with the γ -subunit at the exchangeable binding site 3, with AMP bound, which we have assigned as the weaker of the two sites responsible for mediating AMP/ADP protection against dephosphorylation. The hook makes a lid over the nucleotide-binding site that accounts for a buried surface area of about 500 Å². We obtained crystals of this construct with AMP added to the crystallization mixture, and an AMP molecule is clearly identifiable in the electron density maps at site 3, as well as at the non-exchangeable site. Although we did not see AMP at site 1 in the initial crystals, we subsequently achieved good occupancy at this site by maintaining a higher concentration of AMP during crystallization and handling procedures (data not shown). In contrast, we think that the AMP at site 3 is held in place by the arrangement of the α -hook and that it would have to dissociate before the bound AMP could be released and then exchanged. On the basis of superposition of our earlier structures of the regulatory fragment of AMPK in complex with ADP and Mg-ATP and the structure presented here, we think that the α -hook sequence cannot interact with site 3 when Mg-ATP is bound, mainly

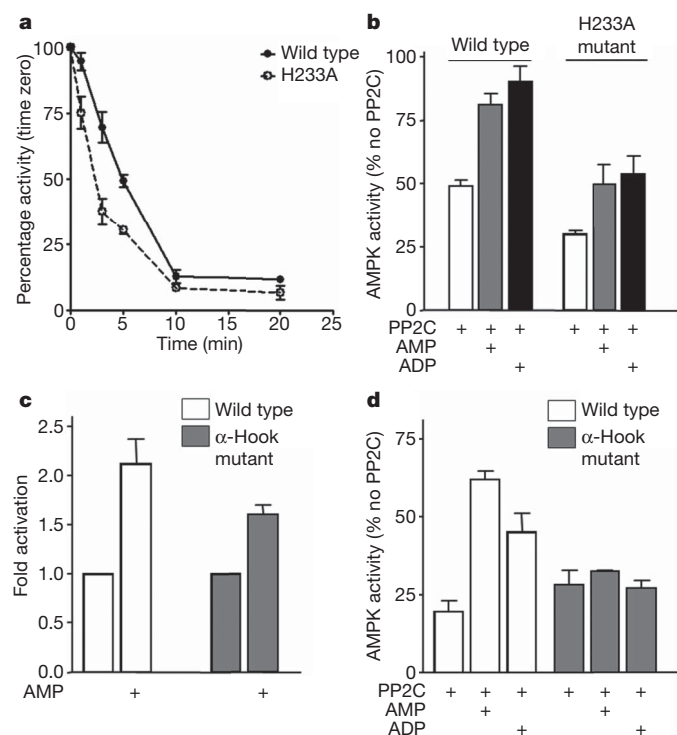


Figure 4 | Mutational analysis of AMPK regulation. **a**, Dephosphorylation rate of the wild type or β -His 233 to alanine kinase domain interface mutant (H233A, corresponding to H235A in β 2). **b**, Protection of dephosphorylation of wild type or H233A mutant by AMP (30 μ M) or ADP (30 μ M) after incubation for 5 min. **c**, Allosteric activation of wild type or α -hook mutant (harbouring mutation of residues R375Q, T377A, D379A and E380A within α 1) by AMP (100 μ M). **d**, AMP (30 μ M) and ADP (30 μ M) protection of wild type or α -hook mutant from dephosphorylation after incubation for 20 min. All results are the mean \pm s.e.m. from at least three independent experiments.

due to the change in conformation of Arg 69 of the γ -subunit, which would generate a steric clash with the α -hook (Supplementary Fig. 11). To test the role of the α -hook in mediating protection against dephosphorylation we generated a mutant in this region (R375Q/T377A/D379A/E380A). The resulting enzyme was allosterically activated by AMP but was not subject to protection against dephosphorylation by AMP or ADP (Fig. 4c, d). Interestingly, the mutation at His 233 described above does retain some protective effect of AMP/ADP (Fig. 4a, b). Given that this mutation would be expected to weaken the interaction between the kinase domain and the regulatory fragment, but not block it, it seems reasonable that AMP/ADP binding would still help to order the α -hook and thus facilitate the recruitment of the kinase domain.

Taking our biochemical and structural data together, we propose the following model for how AMP/ADP, but not Mg-ATP, protects AMPK against dephosphorylation, and thus inactivation (Supplementary Fig. 1). We have demonstrated that the protective effect of AMP/ADP is mediated by its binding to the weaker of the two exchangeable sites, which we have identified as site 3. We have also shown that the α -hook region binds into this site in the presence of AMP and predict that the same situation would occur with ADP. We further suggest that binding of the α -hook acts to restrict the flexibility of the preceding α -linker region (residues 300–370) and, in so doing, promotes the interaction of the kinase domain with the regulatory fragment seen in our crystal structure. This interaction, which mostly involves contacts between the activation loop and the C-terminal domain of β , would therefore act to protect Thr 172 against dephosphorylation. Because the interaction surface of the α -hook with the regulatory fragment is relatively small it is likely that there is a dynamic equilibrium between the α -hook-bound and α -hook-unbound species. If, as our structure indicates, AMP/ADP binding favours the α -hook-bound species but Mg-ATP binding drives formation of the α -hook-unbound species, then the competitive binding of AMP/ADP versus Mg-ATP would control the extent to which the enzyme was protected from dephosphorylation and inactivation.

METHODS SUMMARY

AMPK complexes were expressed in *Escherichia coli* BL21 (DE3) cells, purified by affinity chromatography using nickel-Sepharose and phosphorylated by incubation with CaMKK β as described previously⁹. AMPK activity was determined using 0.2 mM SAMS peptide⁹, 0.2 mM ATP and 5 mM MgCl₂. Dephosphorylation of AMPK by recombinant PP2C- α was monitored either by measuring AMPK activity using the SAMS peptide assay or by western blotting of phospho-T172. Western blot signals for phospho-T172 and total AMPK α -subunit (determined using sheep anti- α 1 or anti- α 2 antibodies) were quantified using the Li-Cor Odyssey infrared imaging system. Uncorrected fluorescence spectra of the nucleotides (3'-(7-diethylaminocoumarin-3-carboxylamino)-3'-deoxy-ADP (C-ADP), and 3'-(7-diethylaminocoumarin-3-carboxylamino)-3'-deoxy-ATP (C-ATP) (both gifts from M. Webb)) and NADH and their complexes were recorded at 20 °C using a Jasco FP-6300 fluorimeter. Binding of nucleotides was monitored by titrating nucleotide (4–10 μ M) with AMPK. Dissociation constants for AMP, ADP and ATP were determined using competition assays. The engineered crystallization construct was expressed as a His-tag fusion protein in *E. coli*. Purified protein was phosphorylated using CaMKK β before mixing with AMP and staurosporine. Crystals were grown by the hanging-drop method using isopropanol and MPD as precipitant. Diffraction data were collected on the Diamond Light Source, Oxford and processed using Denzo and Scalepack²⁸. The structure was solved by molecular replacement using Amore²⁹ and standard refinement was carried out with Refmac³⁰ with manual model building with COOT. Figures were created with Pymol (<http://pymol.sourceforge.net/>).

Received 30 July 2010; accepted 16 February 2011.

Published online 13 March 2011.

- Carling, D. The AMP-activated protein kinase cascade—a unifying system for energy control. *Trends Biochem. Sci.* **29**, 18–24 (2004).
- Hardie, D. G. AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. *Nature Rev. Mol. Cell Biol.* **8**, 774–785 (2007).
- Kahn, B. B., Alquier, T., Carling, D. & Hardie, D. G. AMP-activated protein kinase: ancient energy gauge provides clues to modern understanding of metabolism. *Cell Metab.* **1**, 15–25 (2005).
- Shackelford, D. B. & Shaw, R. J. The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. *Nature Rev. Cancer* **9**, 563–575 (2009).

- Cool, B. *et al.* Identification and characterization of a small molecule AMPK activator that treats key components of type 2 diabetes and the metabolic syndrome. *Cell Metab.* **3**, 403–416 (2006).
- Huang, X. *et al.* Important role of the LKB1-AMPK pathway in suppressing tumorigenesis in PTEN-deficient mice. *Biochem. J.* **412**, 211–221 (2008).
- Stein, S. C., Woods, A., Jones, N. A., Davison, M. D. & Carling, D. The regulation of AMP-activated protein kinase by phosphorylation. *Biochem. J.* **345**, 437–443 (2000).
- Oakhill, J. S. *et al.* β -Subunit myristoylation is the gatekeeper for initiating metabolic stress sensing by AMP-activated protein kinase (AMPK). *Proc. Natl Acad. Sci. USA* **107**, 19237–19241 (2010).
- Sanders, M. J., Grondin, P. O., Hegarty, B. D., Snowden, M. A. & Carling, D. Investigating the mechanism for AMP activation of the AMP-activated protein kinase cascade. *Biochem. J.* **403**, 139–148 (2007).
- Minokoshi, Y. *et al.* Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature* **415**, 339–343 (2002).
- Yamauchi, T. *et al.* Adiponectin stimulates glucose utilization and fatty-acid oxidation by activating AMP-activated protein kinase. *Nature Med.* **8**, 1288–1295 (2002).
- Watt, M. J. *et al.* CNTF reverses obesity-induced insulin resistance by activating skeletal muscle AMPK. *Nature Med.* **12**, 541–548 (2006).
- Andersson, U. *et al.* AMP-activated protein kinase plays a role in the control of food intake. *J. Biol. Chem.* **279**, 12005–12008 (2004).
- Minokoshi, Y. *et al.* AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus. *Nature* **428**, 569–574 (2004).
- Hardie, D. G., Carling, D. & Sim, A. T. R. The AMP-activated protein kinase: a multisubstrate regulator of lipid metabolism. *Trends Biochem. Sci.* **14**, 20–23 (1989).
- Suter, M. *et al.* Dissecting the role of 5'-AMP for allosteric stimulation, activation, and deactivation of AMP-activated protein kinase. *J. Biol. Chem.* **281**, 32207–32216 (2006).
- Carling, D., Sanders, M. J. & Woods, A. The regulation of AMP-activated protein kinase by upstream kinases. *Int. J. Obes.* **32** (Suppl. 4), S55–S59 (2008).
- Momcilovic, M., Hong, S. P. & Carlson, M. Mammalian TAK1 activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase *in vitro*. *J. Biol. Chem.* **281**, 25336–25343 (2006).
- Davies, S. P., Helps, N. R., Cohen, P. T. & Hardie, D. G. 5'-AMP inhibits dephosphorylation, as well as promoting phosphorylation, of the AMP-activated protein kinase. Studies using bacterially expressed human protein phosphatase-2C α and native bovine protein phosphatase-2AC. *FEBS Lett.* **377**, 421–425 (1995).
- Cheung, P. C., Salt, I. P., Davies, S. P., Hardie, D. G. & Carling, D. Characterization of AMP-activated protein kinase γ -subunit isoforms and their role in AMP binding. *Biochem. J.* **346**, 659–669 (2000).
- Xiao, B. *et al.* Structural basis for AMP binding to mammalian AMP-activated protein kinase. *Nature* **449**, 496–500 (2007).
- Kemp, B. E., Oakhill, J. S. & Scott, J. W. AMPK structure and regulation from three angles. *Structure* **15**, 1161–1163 (2007).
- Carling, D., Clarke, P. R., Zammit, V. A. & Hardie, D. G. Purification and characterization of the AMP-activated protein kinase. Copurification of acetyl-CoA carboxylase kinase and 3-hydroxy-3-methylglutaryl-CoA reductase kinase activities. *Eur. J. Biochem.* **186**, 129–136 (1989).
- Veech, R. L., Lawson, J. W., Cornell, N. W. & Krebs, H. A. Cytosolic phosphorylation potential. *J. Biol. Chem.* **254**, 6538–6547 (1979).
- Hellsten, Y., Richter, E. A., Kiens, B. & Bangsbo, J. AMP deamination and purine exchange in human skeletal muscle during and after intense exercise. *J. Physiol.* **520**, 909–920 (1999).
- McConnell, G. K. *et al.* Short-term exercise training in humans reduces AMPK signalling during prolonged exercise independent of muscle glycogen. *J. Physiol. (Lond.)* **568**, 665–676 (2005).
- Little, D. R. *et al.* A conserved mechanism of autoinhibition for the AMPK kinase domain: ATP-binding site and catalytic loop refolding as a means of regulation. *Acta Crystallogr. F* **66**, 143–151 (2010).
- Otwinowski, Z. & Minor, W. Data Collection and Processing. *Proc. CCP4 Study Weekend* 556–562 (SERC, 1993).
- Navaza, J. AMoRe: an Automated Package for Molecular Replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
- CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Webb for gift of coumarin nucleotides, J. Skehel for comments on the manuscript and S. Smerdon for discussion and assistance. Work in both laboratories is supported by the MRC and we gratefully acknowledge Diamond Light Source for synchrotron access.

Author Contributions B.X., M.J.S., E.U., R.H., F.V.M., D. Carmina, C.J., P.A.W., J.F.E., L.F.H., P.S., S.A.H., R.A. and S.R.M. performed experiments. All authors contributed to data analysis, experimental design and manuscript writing.

Author Information Coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 2Y8L, 2Y8Q, 2Y94 and 2YA3. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.C. (david.carling@csc.mrc.ac.uk) or S.J.G. (sgambli@nimr.mrc.ac.uk).

Electromagnetically induced transparency and slow light with optomechanics

A. H. Safavi-Naeini^{1*}, T. P. Mayer Alegre^{1*}, J. Chan¹, M. Eichenfield¹, M. Winger¹, Q. Lin¹, J. T. Hill¹, D. E. Chang^{2,3} & O. Painter¹

Controlling the interaction between localized optical and mechanical excitations has recently become possible following advances in micro- and nanofabrication techniques^{1,2}. So far, most experimental studies of optomechanics have focused on measurement and control of the mechanical subsystem through its interaction with optics, and have led to the experimental demonstration of dynamical back-action cooling and optical rigidity of the mechanical system^{1,3}. Conversely, the optical response of these systems is also modified in the presence of mechanical interactions, leading to effects such as electromagnetically induced transparency⁴ (EIT) and parametric normal-mode splitting⁵. In atomic systems, studies^{6,7} of slow and stopped light (applicable to modern optical networks⁸ and future quantum networks⁹) have thrust EIT to the forefront of experimental study during the past two decades. Here we demonstrate EIT and tunable optical delays in a nanoscale optomechanical crystal, using the optomechanical nonlinearity to control the velocity of light by way of engineered photon–phonon interactions. Our device is fabricated by simply etching holes into a thin film of silicon. At low temperature (8.7 kelvin), we report an optically tunable delay of 50 nanoseconds with near-unity optical transparency, and superluminal light with a 1.4 microsecond signal advance. These results, while indicating significant progress towards an integrated quantum optomechanical memory¹⁰, are also relevant to classical signal processing applications. Measurements at room temperature in the analogous regime of electromagnetically induced absorption show the utility of these chip-scale optomechanical systems for optical buffering, amplification, and filtering of microwave-over-optical signals.

It is by now well known that the optical properties of matter can be dramatically modified by using a secondary light beam, approximately resonant with an internal process of the material system. As an example, an opaque object can be made transparent in the presence of a control beam; this is the phenomenon of EIT. A remarkable feature of EIT is the drastic reduction in the group velocity of light passing through the material, achieved inside a practically lossless transparency window. This aspect of the effect has been used in schemes whereby light may be slowed and stopped, making it an important building block in quantum information and communication proposals, as well as of great practical interest in classical optics and photonics. A simple upper-bound for the storage time in EIT-based proposals is the lifetime related to the internal processes of the material. These lifetimes can be extremely long in atomic gases, with storage times of the order of one second having been demonstrated¹¹ in Bose–Einstein condensates. Part of the vision for future scalable quantum networks has involved extending the remarkable results achieved in atomic experiments to a more readily deployable domain.

In the solid state, EIT has been demonstrated in quantum wells, dots and nitrogen–vacancy centres^{12–14}. But the fast dephasing rates and inhomogeneous broadening of solid-state electronic resonances have led to a plethora of other methods and techniques. Elegant experiments

with stimulated Brillouin scattering in fibres¹⁵ and coherent population oscillations¹⁶ have been used to delay intense classical light. Alternatively, for quantum storage and buffering, techniques related to photon-echo spectroscopy (for example, controllable reversible inhomogeneous broadening¹⁷ and atomic frequency combs¹⁸) have been used successfully to achieve solid-state quantum memories. In chip-scale photonics, dynamically tunable arrays of cavities, displaying EIT, are an intriguing analogue to ensembles of atoms and provide a route to slowing and stopping light all-optically¹⁹. Generally, the elements in the arrays have consisted of coupled optical or plasmonic resonances, and have been demonstrated with couplings engineered to give rise to Fano-like interference²⁰. However, a significant limitation in these all-photon systems is the short optical resonance lifetime. We demonstrate here that in addition to optically controlled switching of a probe beam, as recently presented by others⁴, EIT in an optomechanical cavity may be used to change the group velocity of light significantly. As such, tunable optical delay, with delay times limited by the much longer mechanical resonance lifetime of the optomechanical system, may be achieved. These delays are also attainable across a broad spectrum of wavelengths; indeed, recent circuit cavity electromechanics experiments in the strong-coupling regime have demonstrated EIT and group velocity control at microwave frequencies²¹. Additionally, the ability to create arrays of coupled devices by ‘printing’ optomechanical circuits onto a Si microchip^{22–24}, for example, allows one to create a much larger delay-bandwidth product (scaling as \sqrt{N} , N being the number of cavity elements)¹⁰; such arrays provide a means to sample an incoming optical pulse shape, store it and retrieve it, much like an ensemble of atoms does in atomic EIT.

EIT in optomechanical systems can be understood physically as follows. The conventional radiation pressure interaction between a near-resonant cavity light field and mechanical motion is modelled by the nonlinear Hamiltonian $H_{\text{int}} = \hbar g \hat{a}^\dagger \hat{a} (\hat{b} + \hat{b}^\dagger)$. Here \hat{a} (\hat{a}^\dagger) and \hat{b} (\hat{b}^\dagger) are the annihilation (creation) operators of photon and phonon resonator quanta, respectively, g is the optomechanical coupling rate corresponding physically to the shift in the optical mode’s frequency due to the zero-point fluctuations of the phonon mode, and \hbar is $h/2\pi$, where h is Planck’s constant. By driving the system with an intense red-detuned optical ‘control’ beam at frequency ω_c , as shown in Fig. 1a, the form of the effective interaction changes (in the resolved sideband limit) to that of a beam-splitter-like Hamiltonian, $H_{\text{int}} = \hbar G (\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger)$. Here, the zero-point-motion coupling rate g is replaced by a much stronger parametric coupling rate $G = g\sqrt{\langle n_c \rangle}$ between light and mechanics, where $\langle n_c \rangle$ is the stored intracavity photon number induced by the control beam. Viewed in a dressed-state picture, with the control beam detuning from the optical cavity resonance (ω_o) set equal to the mechanical frequency (ω_m), $\Delta_{\text{OC}} \equiv \omega_o - \omega_c \approx \omega_m$, the optical and mechanical modes \hat{a} and \hat{b} become coupled (denoted \hat{a}_d and \hat{b}_d in Fig. 1b). The dressed mechanical mode, now effectively a phonon–photon polariton, takes on a weakly photonic nature, coupling it to the optical loss channels at a rate $\gamma_{\text{om}} \equiv C\gamma_i$, where the optomechanical

¹Thomas J. Watson Sr Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA. ²Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125, USA. ³Center for the Physics of Information, California Institute of Technology, Pasadena, California 91125, USA.

*These authors contributed equally to this work.

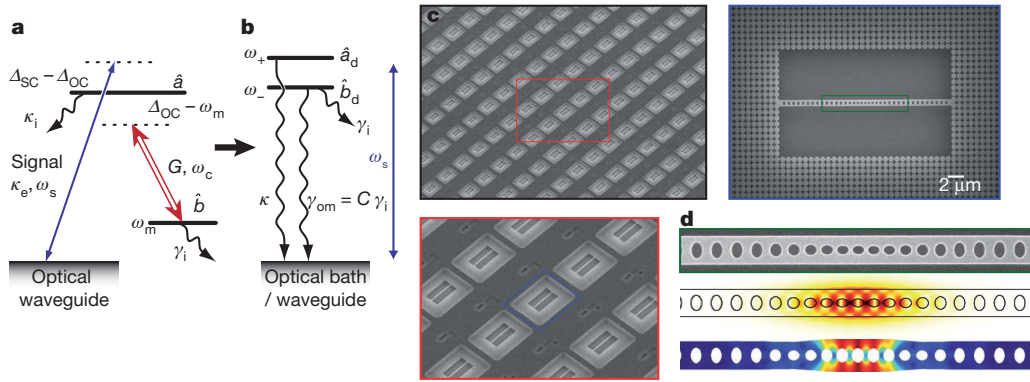


Figure 1 | Optomechanical system. **a**, Level-diagram picture, showing three ‘levels’ that represent the optical mode \hat{a} , the mechanical mode \hat{b} and the ‘bath’ of optical waveguide modes. The transitions between modes driven by the signal and control beams are indicated by blue and red double-headed arrows, respectively. Wavy black arrows indicate decay from the different modes. See text for definitions of symbols in **a** and **b**. **b**, The control beam at ω_c drives the transition between the optical and mechanical mode, dressing the optical and mechanical modes, resulting in the dressed state picture with dressed modes \hat{a}_d

and \hat{b}_d . **c**, Series of scanning electron micrographs, showing large array of optomechanical crystal nanocavities (top-left panel), zoomed-in image of device array (bottom-left panel), and zoomed-in image of top-view of single cavity device (top-right panel). **d**, From top to bottom: scanning electron micrograph of a zoomed-in region showing the OMC defect region; finite-element-method (FEM) simulation results for the optical field showing the electrical field intensity $|\mathbf{E}(\mathbf{r})|$; FEM-simulated mechanical mode with the total displacement $|\mathbf{Q}(\mathbf{r})|$ shown.

cooperativity is defined as $C \equiv 4G^2/\kappa\gamma_i$ for an optical cavity decay rate of κ , and an intrinsic mechanical resonance damping rate of γ_i .

The drive-dependent loss rate γ_{om} has been viewed in most previous studies as an incoherent, quantum-limited loss channel, and was used in recent experiments to cool the mechanical resonator close to its quantum ground state²⁵. In the dressed mode picture, by analogy to the dressed state view of EIT⁷, it becomes clear that a coherent cancellation of the loss channels in the dressed optical and mechanical modes is possible, and can be used to switch the system from absorptive to transmittive in a narrow band around cavity resonance. Much as in atomic EIT, this effect causes an extremely steep dispersion for the transmitted probe photons, with a group delay on resonance of (see Supplementary Information)

$$\tau^{(T)}|_{\omega=\omega_m} = \frac{2}{\gamma_i(1+C)(1-(\kappa_e/\kappa)+C)} \quad (1)$$

where κ_e is the optical coupling rate between the external optical waveguide and the optical cavity, and the delay is dynamically tunable via the control beam intensity through C .

Nano- and micro-optomechanical resonators take a variety of forms, among which optomechanical crystals (OMC) have been used to demonstrate large radiation-pressure-induced interaction strengths between gigahertz mechanical resonances and near-infrared optical resonances²⁴. The nanobeam OMC cavity used in this study (Fig. 1c and d) uses a periodic free-standing Si structure to create high-Q co-localized optical and mechanical resonances. These devices can be printed and etched into the surface of a Si chip in large arrays (Fig. 1c), and are designed to operate optically in the telecommunications band ($\lambda_o = 1,550$ nm) and acoustically at microwave frequencies ($\omega_m/2\pi = 3.75$ GHz). The theoretical optomechanical coupling rate g between co-localized photon and phonon modes is $g/2\pi \approx 800$ kHz. By optimizing the arrangement of holes in the central cavity region of the nanobeam where light and sound are localized, an intrinsic optical decay rate of $\kappa_i/2\pi \approx 290$ MHz is obtained for the optical cavity mode, placing the optomechanical system in the resolved sideband regime ($\omega_m/\kappa_i \gg 1$) necessary for EIT. The corresponding mechanical resonance is measured to have an intrinsic damping rate of $\gamma_i/2\pi \approx 250$ kHz at temperature $T = 8.7$ K, corresponding to a mechanical Q -factor of $Q_m = 15,000$. Light is coupled into and out of the device using a specially prepared optical fibre taper, which when placed in the near-field of the nanobeam cavity couples the guided modes of the taper evanescently to the optical resonances of the nanobeam (see Supplementary Information for details of the optical cavity loading).

In order to characterize the near-resonance optical reflection of the cavity system, a sideband of the control beam is created using electro-optic modulation (see Methods and Supplementary Information), forming a weak signal beam with tunable frequency ω_s . The results of measurements performed at a cryogenic temperature of 8.7 K are shown in Fig. 2. Here, the control beam laser power was varied from 6 μ W ($\langle n_c \rangle = 25$) to nearly 250 μ W ($\langle n_c \rangle = 1,040$). The frequencies of both the control and signal beams are swept in order to map out the system dependence on control beam detuning, Δ_{OC} , and the two-photon detuning, $\Delta_{SC} = \omega_s - \omega_c$. The resulting reflected optical signal intensity, separated from the control beam via a modulation and lock-in technique (see Methods and Supplementary Information), is shown in Fig. 2a for a series of control laser detunings. Visible in each of the plots is a broad resonance corresponding to the bare optical cavity response with loaded linewidth $\kappa/2\pi \approx 900$ MHz. A much narrower reflection dip feature, corresponding to the transparency window, can also be seen near the cavity line centre. The position of the narrow reflection dip tracks with a two-photon detuning equal to the mechanical resonance frequency, $\Delta_{SC} \approx \omega_m$. This region is shown in more detail in Fig. 2b, where the Fano-like structure of the optical response is apparent. Each curve in Fig. 2a and b is a horizontal slice of the data presented in Fig. 2c, where the reflectivity is plotted as a function of both ω_c and Δ_{SC} . The transparency window is shown to be fully controllable via the applied light field, the window expanding and contracting with the control beam laser power (Fig. 2d and e). At the maximum stable control power (unstable regions due to a thermo-optic bistability induced by optical absorption are shown as hatched regions in Fig. 2c), a transparency window approaching 5 MHz is obtained.

Model fits to the reflection spectra (see Supplementary Information) are shown as solid curves in Fig. 2a and b. The resulting fitted values for $\gamma_{om} = 4G^2/\kappa$ for each control power are shown in Fig. 2e. A linear fit to the extracted data yields a value for the zero-point-motion coupling constant of $g/2\pi = 800$ kHz, in agreement with the value obtained from independent optical transduction measurements of the thermal Brownian motion of the mechanical oscillator²⁴. In addition to the intensity response of the optomechanical cavity there is the phase response, which provides a measure of the group delay of the modulated optical signal beam as it passes through the cavity. For the 89-kHz modulation of the signal beam used in our experiments, corresponding to a free-space signal wavelength of ~ 3.4 km, phase shifts between the modulation sidebands and the signal carrier on the order of a fraction of a radian are measured in the region where Δ_{SC} is within a

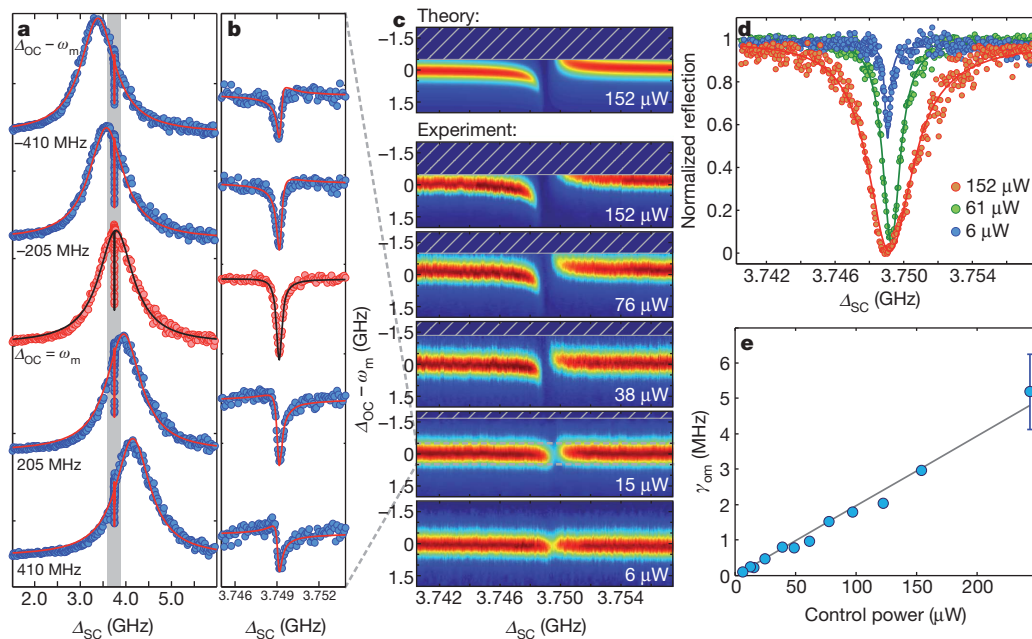


Figure 2 | Optical reflection response at temperature $T = 8.7$ K. **a**, Measured normalized reflection (dots) of the signal beam as a function of the two-photon detuning (Δ_{SC}) for a control beam power of $15 \mu\text{W}$. **b**, Zoom-in of the reflected signal about the transparency window. Each spectrum in **a** and **b** corresponds to a different control laser detuning ($\Delta_{OC} - \omega_m$) as indicated. Solid curves correspond to model fits to the data (see Supplementary Information). **c**, Intensity plots for the signal beam reflection as a function of both control laser detuning (Δ_{OC}) and two-photon detuning (Δ_{SC}) for a series of different

control beam powers (as indicated). The hatched areas are unstable regions for the control laser detuning at the given input power. The top plot is the theoretically predicted reflection spectrum for the highest control beam power. **d**, Transparency window versus control beam power for control laser detuning $\Delta_{OC} \approx \omega_m$. **e**, Transparency window bandwidth ($\gamma_{om} = 4G^2/\kappa$) versus control beam power (error bars indicate the standard deviation in the fit of γ_{om} to the EIT intensity spectra versus Δ_{OC}). The solid line represents the bandwidth scaling for a single best-fit value of $g/2\pi = 800$ kHz.

mechanical linewidth of ω_m . The measured phase-shifts for the reflected signal beam correspond to advances in time of the modulated signal, pointing to causality-preserving superluminal effects. A plot of the peak effective signal advance versus control beam power is plotted in Fig. 3a, ascertained from a fit to the reflection phase response spectra (Fig. 3b). For the highest control power, the reflected signal is advanced by $1.3 \mu\text{s}$, roughly 7,000 times longer than the bare optical cavity lifetime.

The delay in transmission is directly related to the advance on reflection through the bare cavity transmission contrast (measured independently; see Supplementary Information). As such, we plot the corresponding transmission group delay of the signal in Fig. 3c. The theoretical delay/advance of the modulated signal beam for system parameters given by fits to the EIT intensity spectra are shown as dashed curves in Fig. 3a and c, indicating good agreement with the measured phase response. As can be seen in these data, the maximum measured transmission delay is $\tau^{(T)} \approx 50$ ns, which—although corresponding to significant slowing of light (to a velocity of $v_g \approx 40 \text{ m s}^{-1}$) through the few-micrometre-long structure—is much smaller than the measured reflected signal advance or the limit set by the intrinsic mechanical damping ($2/\gamma_i \approx 1.4 \mu\text{s}$). This is due to the weak loading of the optical cavity in these experiments (see Supplementary Information), and the resulting small fraction of transmitted light that actually passes through the cavity.

In addition to the observed EIT-like behaviour of the optomechanical system, a similar phenomena to that of electromagnetically induced absorption (EIA)²⁶ in atomic systems can be realized by setting the detuning of the control beam to the blue side of the optomechanical cavity resonance ($\Delta_{OC} < 0$). Under blue-detuned pumping, the effective Hamiltonian for the optical signal and mechanical phonon mode becomes one of parametric amplification, $H_{\text{int}} = \hbar G(\hat{a}^\dagger \hat{b}^\dagger + \hat{a} \hat{b})$. The measured reflection spectrum from the OMC is shown in Fig. 3d, where the reflectivity of the cavity system is seen to be enhanced

around the two-photon detuning $\Delta_{SC} \approx \omega_m$, a result of the increased ‘absorption’ (feeding) of photons into the cavity. As discussed further in the Supplementary Information, at even higher control beam powers such that $C \gtrsim 1$, the system switches from EIA to parametric amplification, resulting in optical signal amplification, and eventually phonon-lasing.

Reflection spectroscopy at room temperature (296 K) of the optomechanical cavity has also been performed (presented in Supplementary Information), and yields similar results to that of the cryogenic measurements, albeit with a larger value of $\langle n_c \rangle$ required to reach a given cooperativity (see Fig. 3e) owing to the larger intrinsic mechanical dissipation at room temperature ($\gamma_i = 2\pi \times 1.9$ MHz). Beyond the initial demonstrations of EIT and EIA behaviour in the OMC cavities presented here, it is fruitful to consider the bandwidth and signal delay limits that might be attainable with future improvements in device material or geometry. For instance, the transparency bandwidth of the current devices is limited by two-photon absorption of the control beam in the silicon cavities; a move to larger-bandgap dielectric materials, such as silicon nitride, should allow intra-cavity photon numbers of 10^6 (limited by linear material absorption), resulting in a transparency window approaching $G = g\sqrt{\langle n_c \rangle} \approx 2\pi(1 \text{ GHz})$. Also, recent research into low-loss GHz mechanic resonators²⁷ should enable slow light optical delays approaching $10 \mu\text{s}$ at room temperature, roughly equivalent to the optical path length of a kilometre of optical fibre. Much like the acoustic wave devices used in electronic systems²⁸, optomechanical devices with these attributes would enable chip-scale microwave photonic systems capable of advanced signal processing in the optical domain, such as that needed for emerging broadband wireless access networks or more specialized applications, such as true-time delays in radar systems⁸.

The limiting factor for quantum applications of optomechanical systems is the re-thermalization time of the mechanical resonator, $\tau_{\text{th}} = \hbar Q_m/kT$, which in the case of a quantum optical memory

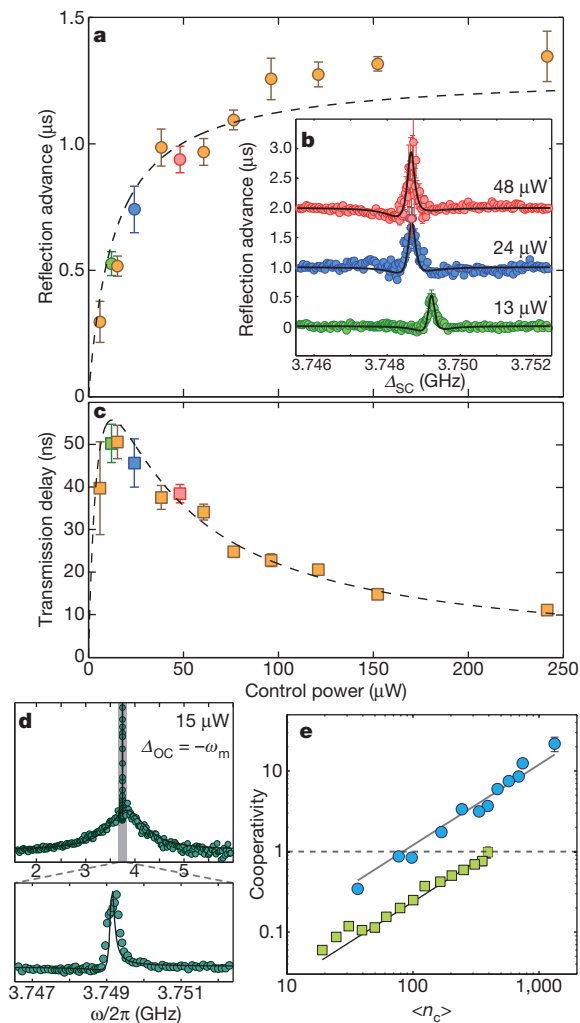


Figure 3 | Measured temporal shifts and amplification. **a**, Maximum measured reflected signal advance as a function of the control beam power. **b**, Measured reflected signal advance versus two-photon detuning, Δ_{SC} . Solid curves correspond to fit from model (see Supplementary Information). Curves at different control powers are shifted for clarity. **c**, Inferred maximum transmitted signal delay versus control beam power. Dashed lines in **a** and **c** are theoretical advance/delay times predicted from model of optomechanical system based on intensity response of the optomechanical system. **d**, Measured signal reflection as a function of two-photon detuning for the control laser blue-detuned from the cavity. **e**, Measured cooperativity for sample temperature of 296 K (squares) and 8.7 K (circles) as a function of the average number of control photons inside the cavity. Error bars indicate the standard deviation in the model fit to the EIT spectral data at each control beam power.

represents the average storage time of a single photon before excitation of the system by a thermal bath phonon. For the devices studied here, despite the optical cooling and reduced phonon occupancy of the mechanical resonator provided by the control beam (the cooling rate being equal to the transparency window bandwidth¹⁰), the re-thermalization time is limited to $\tau_{th} \approx 12$ ns by the 8.7 K bath temperature. Reducing the operating temperature further to a value below 100 mK (routinely attained in a dilution refrigerator) would not only increase the re-thermalization time through a lower bath temperature, but should also result in a significant increase in the mechanical Q -factor. Taken together, the resulting re-thermalization time in the current OMC devices at $T = 100$ mK is likely to be of the order of 100 μs, which although not nearly as long as what has been achieved in atomic systems¹¹, still represents a substantial storage time compared to the realizable GHz bandwidth of the system. Additionally, optomechanical processes similar to the EIT behaviour measured here have also been proposed^{29,30} to provide an optical interface between, for instance,

atomic and superconducting circuit quantum systems, enabling the formation of hybrid quantum networks.

METHODS SUMMARY

Fabrication. The nanobeam cavities were fabricated using a silicon-on-insulator wafer from SOITEC (resistivity $\rho = 4\text{--}20\ \Omega\text{ cm}$, device layer thickness $t = 220$ nm, buried-oxide layer thickness $2\ \mu\text{m}$). The cavity geometry is defined by electron beam lithography followed by inductively-coupled-plasma reactive ion etching to transfer the pattern through the 220-nm silicon device layer. The cavities were then undercut using an HF:H₂O solution to remove the buried oxide layer, and cleaned using a piranha etch/HF etch cycle. The dimensions and design of the nanobeam will be discussed in detail elsewhere.

Experimental set-up. We demonstrate EIT via reflection measurements of the optically pumped system at varying $\langle n_c \rangle$. Using the experimental set-up shown in Supplementary Information, a laser beam at ω_c (the control beam) is sent through an electro-optical modulator with drive frequency equal to Δ_{SC} , creating an optical sideband at frequency ω_s (the signal beam), which is amplitude modulated at $\omega_{LI}/2\pi = 89$ kHz. Since the control beam is detuned from the cavity by $|\Delta_{OC}| \gg \kappa$, it is effectively filtered when looking in reflection, while the modulated signal beam at $\omega_c \pm \Delta_{SC}$ (where the sign is that of Δ_{OC}) is near resonance with the optical cavity and is reflected. The reflected signal beam is detected using a 12-GHz New Focus PIN photo-diode, with the output electrical signal sent to a lock-in amplifier where the component related to the modulated tone ($\omega_{LI} = 89$ kHz) is amplified and sent to an oscilloscope. By scanning both the laser frequency ω_c and the two-photon detuning Δ_{SC} , a full map of the reflectivity is obtained. Additionally, by using a lock-in amplifier, the phase of the modulated signal sidebands relative to the carrier can be measured, giving a direct measurement of the group delay imparted on the optical signal beam by the optomechanical cavity.

Received 8 December 2010; accepted 21 February 2011.

Published online 16 March 2011.

- Kippenberg, T. J. & Vahala, K. J. Cavity optomechanics: back-action at the mesoscale. *Science* **321**, 1172–1176 (2008).
- Favero, I. & Karrai, K. Optomechanics of deformable optical cavities. *Nature Photon.* **3**, 201–205 (2009).
- Braginsky, V. B. *Measurement of Weak Forces in Physics Experiments* (Univ. Chicago Press, 1977).
- Weis, S. *et al.* Optomechanically induced transparency. *Science* **330**, 1520–1523 (2010).
- Gröblacher, S., Hammerer, K., Vanner, M. & Aspelmeyer, M. Observation of strong coupling between a micromechanical resonator and an optical cavity field. *Nature* **460**, 724–727 (2009).
- Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. Light speed reduction to 17 metres per second in an ultracold atomic gas. *Nature* **397**, 594–598 (1999).
- Fleischhauer, M., Imamoglu, A. & Marangos, J. P. Electromagnetically induced transparency: optics in coherent media. *Rev. Mod. Phys.* **77**, 633–673 (2005).
- Boyd, R. W. & Gauthier, D. J. Controlling the velocity of light pulses. *Science* **326**, 1074–1077 (2009).
- Kimble, H. J. The quantum internet. *Nature* **453**, 1023–1030 (2008).
- Chang, D., Safavi-Naeini, A. H., Hafezi, M. & Painter, O. Slowing and stopping light using an optomechanical crystal array. *N. J. Phys.* **13**, 023003 (2011).
- Zhang, R., Garner, S. R. & Hau, L. V. Creation of long-term coherent optical memory via controlled nonlinear interactions in Bose-Einstein condensates. *Phys. Rev. Lett.* **103**, 233602 (2009).
- Phillips, M. C. *et al.* Electromagnetically induced transparency in semiconductors via biexciton coherence. *Phys. Rev. Lett.* **91**, 183602 (2003).
- Santori, C. *et al.* Coherent population trapping of single spins in diamond under optical excitation. *Phys. Rev. Lett.* **97**, 247401 (2006).
- Xu, X. *et al.* Coherent population trapping of an electron spin in a single negatively charged quantum dot. *Nature Phys.* **4**, 692–695 (2008).
- Thévenaz, L. Slow and fast light in optical fibres. *Nature Photon.* **2**, 474–481 (2008).
- Bigelow, M. S., Lepeshkin, N. N. & Boyd, R. W. Superluminal and slow light propagation in a room-temperature solid. *Science* **301**, 200–202 (2003).
- Afzelius, M., Simon, C., de Riedmatten, H. & Gisin, N. Multimode quantum memory based on atomic frequency combs. *Phys. Rev. A* **79**, 052329 (2009).
- de Riedmatten, H., Afzelius, M., Staudt, M. U., Simon, C. & Gisin, N. A solid-state light-matter interface at the single-photon level. *Nature* **456**, 773–777 (2008).
- Yanik, M. F., Suh, W., Wang, Z. & Fan, S. Stopping light in a waveguide with an all-optical analog of electromagnetically induced transparency. *Phys. Rev. Lett.* **93**, 233903 (2004).
- Xu, Q. *et al.* Experimental realization of an on-chip all-optical analogue to electromagnetically induced transparency. *Phys. Rev. Lett.* **96**, 123901 (2006).
- Teufel, J. D. *et al.* Circuit cavity electromechanics in the strong coupling regime. *Nature* doi:10.1038/nature09898 (10 March 2011); preprint at (<http://arXiv.org/abs/1011.3067>) (2010).
- Notomi, M., Kuramochi, E. & Tanabe, T. Large-scale arrays of ultrahigh-Q coupled nanocavities. *Nature Photon.* **2**, 741–747 (2008).

23. Li, M. *et al.* Harnessing optical forces in integrated photonic circuits. *Nature* **456**, 480–484 (2008).
24. Eichenfield, M., Chan, J., Camacho, R. M., Vahala, K. J. & Painter, O. Optomechanical crystals. *Nature* **462**, 78–82 (2009).
25. Rocheleau, T. *et al.* Preparation and detection of a mechanical resonator near the ground state of motion. *Nature* **463**, 72–75 (2010); published online 9 December 2009.
26. Lezama, A., Barreiro, S. & Akulshin, A. M. Electromagnetically induced absorption. *Phys. Rev. A* **59**, 4732–4735 (1999).
27. Nguyen, C. T.-C. MEMS technology for timing and frequency control. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **54**, 251–270 (2007).
28. Lakin, K., Kline, G. & McCarron, K. Development of miniature filters for wireless applications. *IEEE Trans. Microwave Theory Techn.* **43**, 2933–2939 (1995).
29. Stannigel, K., Rabl, P., Sørensen, A. S., Zoller, P. & Lukin, M. D. Optomechanical transducers for long-distance quantum communication. *Phys. Rev. Lett.* **105**, 220501 (2010).
30. Safavi-Naeini, A. H. & Painter, O. Proposal for an optomechanical traveling wave phonon-photon translator. *N. J. Phys.* **13**, 013017 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Schwab for providing the microwave modulation source used in this work. This work was supported by the DARPA/MTO ORCHID programme through a grant from AFOSR, and the Kavli Nanoscience Institute at Caltech. A.H.S.-N. and J.C. acknowledge support from NSERC.

Author Contributions J.C., A.H.S.-N. and M.E. performed the device design, and J.C. performed the device fabrication with support from M.W. and J.T.H. Measurements and data analysis were performed by A.H.S.-N. and T.P.M.A., with support from both D.E.C. and Q.L. and supervision by O.P. All authors contributed to the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to O.P. (opainter@caltech.edu).